

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization  
International Bureau



(43) International Publication Date  
12 December 2002 (12.12.2002)

PCT

(10) International Publication Number  
WO 02/098358 A2

- (51) International Patent Classification: **A61K**
- (21) International Application Number: PCT/US02/17594
- (22) International Filing Date: 4 June 2002 (04.06.2002)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:
- |            |                               |    |
|------------|-------------------------------|----|
| 60/295,917 | 4 June 2001 (04.06.2001)      | US |
| 60/350,666 | 13 November 2001 (13.11.2001) | US |
| 60/368,689 | 29 March 2002 (29.03.2002)    | US |
| 60/372,246 | 12 April 2002 (12.04.2002)    | US |
| 10/160,233 | 31 May 2002 (31.05.2002)      | US |
- (71) Applicant: EOS BIOTECHNOLOGY, INC. [US/US];  
225A Gateway Boulevard, South San Francisco, CA 94080 (US).
- (72) Inventors: AFAR, Daniel, E., H.; 435 Visitation Avenue, Brisbane, CA 94005 (US). AGUS, David; 522 North Crescent Drive, Beverly Hills, CA 90210 (US). MACK, David, H.; 2076 Monterey Avenue, Menlo Park, CA 94025 (US).
- (74) Agents: BASTIAN, Kevin, L. et al.; Townsend and Townsend and Crew LLP, Two Embarcadero Center, Eighth Floor, San Francisco, CA 94111 (US).
- (81) Designated States (national): AE, AG, AI, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZM, ZW.
- (84) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SI, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

**Published:**

— without international search report and to be republished upon receipt of that report

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: METHODS OF DIAGNOSIS AND TREATMENT OF ANDROGEN-DEPENDENT PROSTATE CANCER, PROSTATE CANCER UNDERGOING ANDROGEN-WITHDRAWAL, AND ANDROGEN-INDEPENDENT PROSTATE CANCER

(57) Abstract: Described herein are genes whose expression are up regulated or down regulated in prostate cancer. Also described are such genes whose expression is further up-regulated or down-regulated in drug-resistant prostate cancer cells. Related methods and compositions that can be used for diagnosis and treatment of prostate cancer are disclosed. Also described herein are methods that can be used to identify modulators of prostate cancer.



WO 02/098358 A2

WO 02/098358

PCT/US02/17594

METHODS OF DIAGNOSIS AND TREATMENT OF ANDROGEN-DEPENDENT  
PROSTATE CANCER, PROSTATE CANCER UNDERGOING ANDROGEN  
WITHDRAWAL, AND ANDROGEN-INDEPENDENT PROSTATE CANCER

5

CROSS-REFERENCES TO RELATED APPLICATIONS

This application claims priority from the following applications: USSN 60/295,917,  
filed June 4, 2001, USSN 60/368,689, filed March 29, 2002; USSN 60/350,666, filed  
November 13, 2001; and USSN 60/372,246, filed April 12, 2002; each of which is

10 incorporated herein by reference in its entirety.

FIELD OF THE INVENTION

The invention relates to the identification of nucleic acid and protein expression  
profiles and nucleic acids, products, and antibodies thereto that are involved in prostate  
15 cancer; and to the use of such expression profiles and compositions in the diagnosis,  
prognosis, and therapy of prostate cancer. The invention further relates to methods for  
identifying and using agents and/or targets that inhibit prostate cancer.

BACKGROUND OF THE INVENTION

20 Prostate cancer is the most frequently diagnosed cancer and the second leading cause  
of male cancer death in North America and northern Europe. Early detection of prostate  
cancer using a serum test for prostate-specific antigen (PSA) has dramatically improved the  
treatment of the disease (Oesterling (1992) *J. Am. Med. Assoc.* 267:2236-2238). Treatment  
of prostate cancer consists largely of surgical prostatectomy, radiation therapy, androgen  
25 ablation therapy and chemotherapy. Although many prostate cancer patients are effectively  
treated, the current therapies can all induce serious side effects which diminish quality of life.  
Patients who present with metastatic disease are most often treated with androgen-ablation  
therapy. Hormone blockade results in significant regression of the tumor. However, this  
treatment rarely cures the patient and invariably results in progression to androgen-

independent disease, which is incurable. Afrin and Stuart (1994) J.S.C. Med. Assoc. 90:231-236.

The identification of novel therapeutic targets and diagnostic markers is essential for improving the current treatment of prostate cancer patients. Recent advances in molecular medicine have increased the interest in tumor-specific cell surface antigens that could serve as targets for various immunotherapeutic or small molecule strategies. Antigens suitable for immunotherapeutic strategies should be highly expressed in cancer tissues and ideally not expressed in normal adult tissues. Expression in tissues that are dispensable for life, however, may be tolerated. Examples of such antigens include Her2/neu and the B-cell antigen CD20. Humanized monoclonal antibodies directed to Her2/neu (Herceptin) are currently in use for the treatment of metastatic breast cancer. Ross and Fletcher (1998) Stem Cells 16:413-428. Similarly, anti-CD20 monoclonal antibodies (Rituxin) are used to effectively treat non-Hodgkin's lymphoma. Maloney, et al. (1997) Blood 90:2188-2195; Legot and Czuczman (1998) Curr. Opin. Oncol. 10:548-551.

Several potential immunotherapeutic targets have been identified for prostate cancer. They include prostate-specific membrane antigen (PSMA) (Israeli, et al. (1993) Cancer Res. 53:227-230), prostate stem cell antigen (PSCA; Reiter, et al. (1998) Proc. Natl. Acad. Sci. USA 95:1735-1740), and serpentine transmembrane epithelial antigen of the prostate (STEAP; Hubert, et al. (1999) Proc. Natl. Acad. Sci. USA 96:14529-14534). PSMA is a type II transmembrane hydrolase with significant homology to a rat neuropeptidase (Carter, et al. (1996) Proc. Natl. Acad. Sci. USA 93:749-753). Antibodies directed towards PSMA are currently being used to detect metastasized prostate cancer as the Prostascint Scan (Sodec, et al. (1996) Clin. Nucl. Med. 21:759-767) and are also being evaluated for treatment of advanced disease (Gregorakis, et al. (1998) Semin. Urol. Oncol. 16:2-12; Liu, et al. (1998) Cancer Res. 58:4055-4060; Murphy, et al. (1998) J. Urol. 160:2396-2401). In a study on bone metastasis of prostate cancer, only 8 out of 18 patient samples expressed PSMA (Silver, et al. (1997) Clin. Cancer Res. 3:81-85). Therefore, it is clear that other targets need to be identified to manage metastasized disease. PSCA is a member of the Thy-1/Ly-6 family of glycosylphosphatidylinositol-linked plasma membrane proteins (Reiter, et al. (1998) Proc. Natl. Acad. Sci. USA 95:1735-1740). Immunohistochemical data shows that PSCA is up-regulated in the majority of prostate cancer epithelia and is also detected in bone metastasis (Gu, et al. (2000) Oncogene 19:1288-1296). Recent work shows that antibodies directed to

PSCA can prevent metastatic spread of prostate cancer in a mouse model (Saffran, et al. (2001) *Proc. Natl. Acad. Sci. USA* 98:2658-2663). STEAP is a multi-transmembrane prostate-specific protein that may function as a channel or transporter protein (Hubert, et al. (1999) *Proc. Natl. Acad. Sci. USA* 96:14529-14534). Its protein expression is specific to the basolateral membranes of normal prostate and prostate cancer epithelia. STEAP expression was most highly concentrated at cell-cell boundaries, implying a potential function in intercellular communication. Therapeutic monoclonal antibodies have so far not been reported for STEAP.

### SUMMARY OF THE INVENTION

The present invention therefore provides nucleotide sequences of genes that are up- and down-regulated in androgen-independent prostate cancer cells or prostate cells undergoing androgen withdrawal. Such genes are useful for diagnostic purposes, and also as targets for screening for therapeutic compounds that modulate prostate cancer, such as hormones or antibodies. Other aspects of the invention will become apparent to the skilled artisan by the following description of the invention.

In one aspect, the present invention provides a method of detecting an androgen independent prostate cancer-associated transcript in a cell from a patient, the method comprising contacting a biological sample from the patient with a polynucleotide that selectively hybridizes to nucleic acid molecule comprising a sequence at least 80% identical to a sequence as shown in Tables 1A-4.

In one embodiment, the present invention provides a method of determining the level of a prostate cancer associated transcript in a cell from a patient.

In one embodiment, the present invention provides a method of detecting a prostate cancer-associated transcript in a cell from a patient, the method comprising contacting a biological sample from the patient with a polynucleotide that selectively hybridizes to a sequence at least 80% identical to a sequence as shown in Tables 1A-4.

In various embodiments, the polynucleotide selectively hybridizes to a sequence at least 95% identical to a sequence as shown in Tables 1A-4; the polynucleotide comprises a sequence as shown in Tables 1A-4; the biological sample is a tissue sample; the biological sample comprises isolated nucleic acids, e.g., mRNA; the polynucleotide is labeled, e.g., with a fluorescent label; the polynucleotide is immobilized on a solid surface; the patient is



undergoing a therapeutic regimen to treat prostate cancer; the patient is suspected of having metastatic prostate cancer; the patient is a human; the patient is suspected of having a taxol-resistant cancer; or the prostate cancer associated transcript is mRNA.

In other embodiments, the method further comprises the step of amplifying nucleic acid before the step of contacting the biological sample with the polynucleotide.

In another aspect, the present invention provides a method of monitoring the efficacy of a therapeutic treatment of prostate cancer, the method comprising the steps of: (i) providing a biological sample from a patient undergoing the therapeutic treatment; and (ii) determining the level of a prostate cancer-associated transcript in the biological sample by contacting the biological sample with a polynucleotide that selectively hybridizes to a sequence at least 80% identical to a sequence as shown in Tables 1A-4, thereby monitoring the efficacy of the therapy. In a further embodiment, the patient has metastatic prostate cancer. In a further embodiment, the patient has a drug resistant (e.g., taxol resistant) form of prostate cancer.

In one embodiment, the method further comprises the step of: (iii) comparing the level of the prostate cancer-associated transcript to a level of the prostate cancer-associated transcript in a biological sample from the patient prior to, or earlier in, the therapeutic treatment.

Additionally, provided herein is a method of evaluating the effect of a candidate prostate cancer drug comprising administering the drug to a patient and removing a cell sample from the patient. The expression profile of the cell is then determined. This method may further comprise comparing the expression profile to an expression profile of a healthy individual. In a preferred embodiment, said expression profile includes a gene of Tables 1A-4.

In one aspect, the present invention provides an isolated nucleic acid molecule consisting of a polynucleotide sequence as shown in Tables 1A-4.

In one embodiment, an expression vector or cell comprises the isolated nucleic acid.

In one aspect, the present invention provides an isolated polypeptide which is encoded by a nucleic acid molecule having polynucleotide sequence as shown in Tables 1A-4.

In another aspect, the present invention provides an antibody that specifically binds to an isolated polypeptide which is encoded by a nucleic acid molecule having polynucleotide sequence as shown in Tables 1A-4.

In certain embodiments, the antibody is conjugated to an effector component, e.g., a fluorescent label, a radioisotope or a cytotoxic chemical; the antibody is an antibody fragment; or the antibody is humanized.

In one aspect, the present invention provides a method of detecting a prostate cancer cell in a biological sample from a patient, the method comprising contacting the biological sample with an antibody as described herein.

In another aspect, the present invention provides a method of detecting antibodies specific to prostate cancer in a patient, the method comprising contacting a biological sample from the patient with a polypeptide encoded by a nucleic acid comprising a sequence from Tables 1A-4.

In another aspect, the present invention provides a method for identifying a compound that modulates a prostate cancer-associated polypeptide, the method comprising the steps of: a) contacting the compound with a prostate cancer-associated polypeptide, the polypeptide encoded by a polynucleotide that selectively hybridizes to a sequence at least 80% identical to a sequence as shown in Tables 1A-4; and b) determining the functional effect of the compound upon the polypeptide.

In one embodiment, the functional effect is a physical effect, an enzymatic effect, or a chemical effect.

In one embodiment, the polypeptide is expressed in a eukaryotic host cell or cell membrane. In another embodiment, the polypeptide is recombinant.

In one embodiment, the functional effect is determined by measuring ligand binding to the polypeptide.

In another aspect, the present invention provides a method of inhibiting proliferation of a prostate cancer-associated cell to treat prostate cancer in a patient, the method comprising the step of administering to the subject a therapeutically effective amount of a compound identified as described herein.

In one embodiment, the compound is an antibody.

In another aspect, the present invention provides a drug screening assay comprising the steps of: a) administering a test compound to a mammal having prostate cancer or to a cell sample isolated therefrom; b) comparing the level of gene expression of a polynucleotide that selectively hybridizes to a sequence at least 80% identical to a sequence as shown in Tables 1A-4 in a treated cell or mammal with the level of gene expression of the

polynucleotide in a control cell sample or mammal, wherein a test compound that modulates the level of expression of the polynucleotide is a candidate for the treatment of prostate cancer.

In one embodiment, the control is a mammal with prostate cancer or a cell sample therefrom that has not been treated with the test compound. In another embodiment, the control is a normal cell or mammal.

In one embodiment, the test compound is administered in varying amounts or concentrations. In another embodiment, the test compound is administered for varying time periods. In another embodiment, the comparison can occur after addition or removal of the drug candidate.

In one embodiment, the levels of a plurality of polynucleotides that selectively hybridize to a sequence at least 80% identical to a sequence as shown in Tables 1A-4 are individually compared to their respective levels in a control cell sample or mammal. In a preferred embodiment the plurality of polynucleotides is from three to ten.

In another aspect, the present invention provides a method for treating a mammal having prostate cancer comprising administering a compound identified by the assay described herein.

In another aspect, the present invention provides a pharmaceutical composition for treating a mammal having prostate cancer, the composition comprising a compound identified by the assay described herein and a physiologically acceptable excipient.

In one aspect, the present invention provides a method of screening drug candidates by providing a cell expressing a gene that is up- and down-regulated as in a prostate cancer. In one embodiment, a gene is selected from Tables 1A-4. The method further includes adding a drug candidate to the cell and determining the effect of the drug candidate on the expression of the expression profile gene.

In one embodiment, the method of screening drug candidates includes comparing the level of expression in the absence of the drug candidate to the level of expression in the presence of the drug candidate, wherein the concentration of the drug candidate can vary when present, and wherein the comparison can occur after addition or removal of the drug candidate. In a preferred embodiment, the cell expresses at least two expression profile genes. The profile genes may show an increase or decrease.

Also provided is a method of evaluating the effect of a candidate prostate cancer drug comprising administering the drug to a transgenic animal expressing or over-expressing the prostate cancer modulatory protein, or an animal lacking the prostate cancer modulatory protein, for example as a result of a gene knockout.

5 Moreover, provided herein is a biochip comprising one or more nucleic acid segments of Tables 1A-4, wherein the biochip comprises fewer than 1000 nucleic acid probes. Preferably, at least two nucleic acid segments are included. More preferably, at least three nucleic acid segments are included.

10 Furthermore, a method of diagnosing a disorder associated with prostate cancer is provided. The method comprises determining the expression of a gene of Tables 1A-4, in a first tissue type of a first individual, and comparing the distribution to the expression of the gene from a second normal tissue type from the first individual or a second unaffected individual. A difference in the expression indicates that the first individual has a disorder associated with prostate cancer.

15 In a further embodiment, the biochip also includes a polynucleotide sequence of a gene that is not up- and down-regulated in prostate cancer.

In one embodiment a method for screening for a bioactive agent capable of interfering with the binding of a prostate cancer modulating protein (prostate cancer modulatory protein) or a fragment thereof and an antibody which binds to said prostate cancer modulatory protein or fragment thereof. In a preferred embodiment, the method comprises combining a prostate cancer modulatory protein or fragment thereof, a candidate bioactive agent and an antibody which binds to said prostate cancer modulatory protein or fragment thereof. The method further includes determining the binding of said prostate cancer modulatory protein or fragment thereof and said antibody. Wherein there is a change in binding, an agent is identified as an interfering agent. The interfering agent can be an agonist or an antagonist. Preferably, the agent inhibits prostate cancer.

Also provided herein are methods of eliciting an immune response in an individual. In one embodiment a method provided herein comprises administering to an individual a composition comprising a prostate cancer modulating protein, or a fragment thereof. In another embodiment, the protein is encoded by a nucleic acid selected from those of Tables 1A-4.

Further provided herein are compositions capable of eliciting an immune response in an individual. In one embodiment, a composition provided herein comprises a prostate cancer modulating protein, preferably encoded by a nucleic acid of Tables 1A-4, or a fragment thereof, and a pharmaceutically acceptable carrier. In another embodiment, said composition comprises a nucleic acid comprising a sequence encoding a prostate cancer modulating protein, preferably selected from the nucleic acids of Tables 1A-4 and a pharmaceutically acceptable carrier.

Also provided are methods of neutralizing the effect of a prostate cancer protein, or a fragment thereof, comprising contacting an agent specific for said protein with said protein in an amount sufficient to effect neutralization. In another embodiment, the protein is encoded by a nucleic acid selected from those of Tables 1A-4. In another aspect of the invention, a method of treating an individual for prostate cancer is provided. In one embodiment, the method comprises administering to said individual an inhibitor of a prostate cancer modulating protein. In another embodiment, the method comprises administering to a patient having prostate cancer an antibody to a prostate cancer modulating protein conjugated to a therapeutic moiety. Such a therapeutic moiety can be a cytotoxic agent or a radioisotope.

# DETAILED DESCRIPTION OF THE INVENTION

In accordance with the objects outlined above, the present invention provides novel methods for diagnosis and evaluation of androgen-dependent prostate cells (malignant or non-malignant), prostate cells undergoing androgen withdrawal, and androgen-independent prostate cancer, as well as methods for treating androgen-dependent prostate cells (malignant or non-malignant), prostate cancer undergoing androgen withdrawal, and androgen-independent prostate cancer. The current Specification incorporates the text of USSN 09/976,858, filed October 12, 2001, USSN 60/295,917, filed June 4, 2001, USSN 60/368,689, filed March 29, 2002; USSN 60/350,666, filed November 13, 2001; and USSN 60/372,246, filed April 12, 2002.

Table 1A provides unigene cluster identification numbers for the nucleotide sequence of genes that exhibit increased or decreased expression in androgen-independent prostate cancer samples. Table 1A also provides an exemplar accession number that provides a nucleotide sequence that is part of the unigene cluster. The expression patterns of the genes of Table 1A can be broadly defined into the following categories:

Genes that are expressed early in the time course, then drop off in expression, and then express again with emergence of androgen-independence (hi-lo-hi pattern in table 1A). Genes that are expressed early in the time course, then drop off in expression, and do not express again with emergence of androgen-independence (hi-lo-lo pattern in 1A). Genes that are not expressed early in the time course, but express only with emergence of androgen-independence (lo-lo-hi pattern in table 1A). Genes that are not expressed early in the time course, but then express as androgen is withdrawn and continue to express with emergence of androgen-independence (lo-hi-hi pattern in table 1A). Genes that are not expressed early in the time course, but then express as androgen is withdrawn and drop off again with emergence of androgen-independence (lo-hi-lo pattern in table 1A).

Tables 2A-C provide unigene cluster identification numbers for the nucleotide sequence of genes that exhibit increased or decreased expression in androgen-dependent prostate cancer, prostate cancer undergoing androgen withdrawal and androgen-independent prostate cancer. Tables 2A-C also provide an exemplar accession number that provides a nucleotide sequence that is part of the unigene cluster. The expression patterns of the genes of Tables 2A-C can be broadly defined into the following 6 categories:

Genes that are expressed early in the time course of androgen withdrawal, then drop off in expression, and then express again with emergence of androgen-independence (hi-lo-lo-hi pattern in Table 2A). Genes that are expressed early in the time course, then drop off in expression immediately after androgen-withdrawal, and do not express again with emergence of androgen-independence (hi-lo-lo-lo pattern in Table 2A). Genes that are expressed early in the time course, then drop off in expression after several days of androgen withdrawal, and do not express again with emergence of androgen-independence (hi-hi-lo-lo pattern in Table 2A). Genes that are not expressed early in the time course, but express only with emergence of androgen-independence (lo-lo-lo-hi pattern in Table 2A). Genes that are not expressed early in the time course, but then express as androgen is withdrawn and continue to express with emergence of androgen-independence (lo-lo-hi-hi pattern in Table 2A). Genes that are not expressed early in the time course, but then express as androgen is withdrawn and drop off again with emergence of androgen-independence (lo-lo-hi-lo pattern in Table 2A).

## Definitions

The term "androgen ablation therapy" refers to techniques for the removal or destruction of sources of male hormones, such as testosterone. These techniques include, for example, 1) surgical removal of the testicles, 2) medications such as gonadotropin releasing hormone analogs that inhibit testosterone production, or 3) anti-androgenic drugs that block androgen receptors.

The term "androgen-independent prostate cancer protein" or "androgen-independent prostate cancer polynucleotide" or "androgen-independent prostate cancer-associated transcript" refers to nucleic acid and polypeptide polymorphic variants, alleles, mutants, and interspecies homologues that: (1) have a nucleotide sequence that has greater than about 60% nucleotide sequence identity, 65%, 70%, 75%, 80%, 85%, 90%, preferably 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% or greater nucleotide sequence identity, preferably over a region of over a region of at least about 25, 50, 100, 200, 500, 1000, or more nucleotides, to a nucleotide sequence of or associated with a unigene cluster of Tables 1A-4; (2) bind to antibodies, e.g., polyclonal antibodies, raised against an immunogen comprising an amino acid sequence encoded by a nucleotide sequence of or associated with a unigene cluster of Tables 1A-4 and conservatively modified variants thereof; (3) specifically hybridize under stringent hybridization conditions to a nucleic acid sequence, or the complement thereof of Tables 1A-4 and conservatively modified variants thereof; or (4) have an amino acid sequence that has greater than about 60% amino acid sequence identity, 65%, 70%, 75%, 80%, 85%, 90%, preferably 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% or greater amino sequence identity, preferably over a region of over a region of at least about 25, 50, 100, 200, 500, 1000, or more amino acid, to an amino acid sequence encoded by a nucleotide sequence of or associated with a unigene cluster of Tables 1A-4. These polynucleotides or proteins may also be expressed during a period following androgen withdrawal. A polynucleotide or polypeptide sequence is typically from a mammal including, but not limited to, primate, e.g., human; rodent, e.g., rat, mouse, hamster; cow, pig, horse, sheep, or other mammal. A "prostate cancer polypeptide" and a "prostate cancer polynucleotide," include both naturally occurring or recombinant forms, and may refer to those polypeptides or polynucleotides which are expressed in prostate proliferative cells.

A "full length" prostate cancer protein or nucleic acid refers to a prostate cancer polypeptide or polynucleotide sequence, or a variant thereof, that contains the elements normally contained in one or more naturally occurring, wild type prostate cancer

polynucleotide or polypeptide sequences. The “full length” may be prior to, or after, various stages of post-translation processing or splicing, including alternative splicing.

“Biological sample” as used herein is a sample of biological tissue or fluid that contains nucleic acids or polypeptides, e.g., of a prostate cancer protein, polynucleotide or transcript. Such samples include, but are not limited to, tissue isolated from primates, e.g., humans, or rodents, e.g., mice, and rats. Biological samples may also include sections of tissues such as biopsy and autopsy samples, frozen sections taken for histology purposes, blood, plasma, serum, sputum, stool, tears, mucus, hair, skin, etc. Biological samples also include explants and primary and/or transformed cell cultures derived from patient tissues. A biological sample is typically obtained from a eukaryotic organism, most preferably a mammal such as a primate e.g., chimpanzee or human; cow; dog; cat; a rodent, e.g., guinea pig, rat, mouse; rabbit; or a bird; reptile; or fish.

“Providing a biological sample” means to obtain a biological sample for use in methods described in this invention. Most often, this will be done by removing a sample of cells from an animal, but can also be accomplished by using previously isolated cells (e.g., isolated by another person, at another time, and/or for another purpose), by collecting a sample which contains a soluble polypeptide or nucleic acid derived from a prostate cell, or by performing the methods of the invention in vivo. Archival tissues, having treatment or outcome history, will be particularly useful.

The terms “identical” or percent “identity,” in the context of two or more nucleic acids or polypeptide sequences, refer to two or more sequences or subsequences that are the same or have a specified percentage of amino acid residues or nucleotides that are the same (i.e., about 60% identity, preferably 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or higher identity over a specified region, when compared and aligned for maximum correspondence over a comparison window or designated region) as measured using a BLAST or BLAST 2.0 sequence comparison algorithms with default parameters described below, or by manual alignment and visual inspection (see, e.g., NCBI web site <http://www.ncbi.nlm.nih.gov/BLAST/> or the like). Such sequences are then said to be “substantially identical.” This definition also refers to, or may be applied to, the complement of a test sequence. The definition also includes sequences that have deletions and/or additions, as well as those that have substitutions, as well as naturally occurring, e.g., polymorphic or allelic variants, and man-made variants. As described below, the preferred



algorithms can account for gaps and the like. Preferably, identity exists over a region that is at least about 25 amino acids or nucleotides in length, or more preferably over a region that is 50-100 amino acids or nucleotides in length.

For sequence comparison, typically one sequence acts as a reference sequence, to which test sequences are compared. When using a sequence comparison algorithm, test and reference sequences are entered into a computer, subsequence coordinates are designated, if necessary, and sequence algorithm program parameters are designated. Preferably, default program parameters can be used, or alternative parameters can be designated. The sequence comparison algorithm then calculates the percent sequence identities for the test sequences relative to the reference sequence, based on the program parameters.

A "comparison window", as used herein, includes reference to a segment of one of the number of contiguous positions selected from the group consisting typically of from 20 to 600, usually about 50 to about 200, more usually about 100 to about 150 in which a sequence may be compared to a reference sequence of the same number of contiguous positions after the two sequences are optimally aligned. Methods of alignment of sequences for comparison are well-known in the art. Optimal alignment of sequences for comparison can be conducted, e.g., by the local homology algorithm of Smith and Waterman (1981) Appl. Math. 2:482, by the homology alignment algorithm of Needleman and Wunsch (1970) J. Mol. Biol. 48:443-453, by the search for similarity method of Pearson and Lipman (1988) Proc. Nat'l. Acad. Sci. USA 85:2444-2448, by computerized implementations of these algorithms (GAP, BESTFIT, FASTA, and TFASTA in the Wisconsin Genetics Software Package, Genetics Computer Group, 575 Science Dr., Madison, WI), or by manual alignment and visual inspection (see, e.g., Ausubel, et al. (eds. 1995 and supplements) Current Protocols in Molecular Biology Lippincott).

Preferred examples of algorithms that are suitable for determining percent sequence identity and sequence similarity include the BLAST and BLAST 2.0 algorithms, which are described in Altschul, et al. (1977) Nuc. Acids Res. 25:3389-3402 and Altschul, et al. (1990) J. Mol. Biol. 215:403-410. BLAST and BLAST 2.0 are used, with the parameters described herein, to determine percent sequence identity for the nucleic acids and proteins of the invention. Software for performing BLAST analyses is publicly available through the National Center for Biotechnology Information (<http://www.ncbi.nlm.nih.gov/>). This algorithm involves first identifying high scoring sequence pairs (HSPs) by identifying short

words of length W in the query sequence, which either match or satisfy some positive-valued threshold score T when aligned with a word of the same length in a database sequence. T is referred to as the neighborhood word score threshold (Altschul, et al., supra). These initial neighborhood word hits act as seeds for initiating searches to find longer HSPs containing them. The word hits are extended in both directions along each sequence for as far as the cumulative alignment score can be increased. Cumulative scores are calculated using, e.g., for nucleotide sequences, the parameters M (reward score for a pair of matching residues; always > 0) and N (penalty score for mismatching residues; always < 0). For amino acid sequences, a scoring matrix is used to calculate the cumulative score. Extension of the word hits in each direction are halted when: the cumulative alignment score falls off by the quantity X from its maximum achieved value; the cumulative score goes to zero or below, due to the accumulation of one or more negative-scoring residue alignments; or the end of either sequence is reached. The BLAST algorithm parameters W, T, and X determine the sensitivity and speed of the alignment. The BLASTN program (for nucleotide sequences) uses as defaults a wordlength (W) of 11, an expectation (E) of 10, M=5, N=-4 and a comparison of both strands. For amino acid sequences, the BLASTP program uses as defaults a wordlength of 3, and expectation (E) of 10, and the BLOSUM62 scoring matrix (see Henikoff and Henikoff (1989) Proc. Natl. Acad. Sci. USA 89:10915-919) alignments (B) of 50, expectation (E) of 10, M=5, N=-4, and a comparison of both strands.

The BLAST algorithm also performs a statistical analysis of the similarity between two sequences (see, e.g., Karlin and Altschul (1993) Proc. Nat'l. Acad. Sci. USA 90:5873-5787). One measure of similarity provided by the BLAST algorithm is the smallest sum probability (P(N)), which provides an indication of the probability by which a match between two nucleotide or amino acid sequences would occur by chance. For example, a nucleic acid is considered similar to a reference sequence if the smallest sum probability in a comparison of the test nucleic acid to the reference nucleic acid is less than about 0.2, more preferably less than about 0.01, and most preferably less than about 0.001. Log values may be large negative numbers, e.g., 5, 10, 20, 30, 40, 40, 70, 90, 110, 150, 170, etc.

An indication that two nucleic acid sequences or polypeptides are substantially identical is that the polypeptide encoded by the first nucleic acid is immunologically cross reactive with the antibodies raised against the polypeptide encoded by the second nucleic acid, as described below. Thus, a polypeptide is typically substantially identical to a second

polypeptide, e.g., where the two peptides differ only by conservative substitutions. Another indication that two nucleic acid sequences are substantially identical is that the two molecules or their complements hybridize to each other under stringent conditions, as described below. Yet another indication that two nucleic acid sequences are substantially identical is that the same primers can be used to amplify the sequences.

A "host cell" is a naturally occurring cell or a transformed cell that contains an expression vector and supports the replication or expression of the expression vector. Host cells may be cultured cells, explants, cells in vivo, and the like. Host cells may be prokaryotic cells such as *E. coli*, or eukaryotic cells such as yeast, insect, amphibian, or mammalian cells such as CHO, HeLa, and the like (see, e.g., the American Type Culture Collection catalog or web site, [www.atcc.org](http://www.atcc.org)).

The terms "isolated," "purified," or "biologically pure" refer to material that is substantially or essentially free from components that normally accompany it as found in its native state. Purity and homogeneity are typically determined using analytical chemistry techniques such as polyacrylamide gel electrophoresis or high performance liquid chromatography. A protein or nucleic acid that is the predominant species present in a preparation is substantially purified. In particular, an isolated nucleic acid is separated from some open reading frames that naturally flank the gene and encode proteins other than protein encoded by the gene. The term "purified" in some embodiments denotes that a nucleic acid or protein gives rise to essentially one band in an electrophoretic gel. Preferably, it means that the nucleic acid or protein is at least 85% pure, more preferably at least 95% pure, and most preferably at least 99% pure. "Purify" or "purification" in other embodiments means removing at least one contaminant from the composition to be purified. In this sense, purification does not require that the purified compound be homogenous, e.g., 100% pure.

The terms "polypeptide," "peptide," and "protein" are used interchangeably herein to refer to a polymer of amino acid residues. The terms apply to amino acid polymers in which one or more amino acid residue is an artificial chemical mimetic of a corresponding naturally occurring amino acid, as well as to naturally occurring amino acid polymers, those containing modified residues, and non-naturally occurring amino acid polymer. Certain diagnostic methods may evaluate secreted or breakdown products present only because the producing cell is present, and would otherwise be absent in a normal individual.

The term "amino acid" refers to naturally occurring and synthetic amino acids, as well as amino acid analogs and amino acid mimetics that function similarly to the naturally occurring amino acids. Naturally occurring amino acids are those encoded by the genetic code, as well as those amino acids that are later modified, e.g., hydroxyproline,  $\gamma$ -carboxyglutamate, and O-phosphoserine. Amino acid analogs refers to compounds that have the same basic chemical structure as a naturally occurring amino acid, e.g., an  $\alpha$  carbon that is bound to a hydrogen, a carboxyl group, an amino group, and an R group, e.g., homoserine, norleucine, methionine sulfoxide, methionine methyl sulfonium. Such analogs may have modified R groups (e.g., norleucine) or modified peptide backbones, but retain the same basic chemical structure as a naturally occurring amino acid. Amino acid mimetics refers to chemical compounds that have a structure that is different from the general chemical structure of an amino acid, but that functions similarly to a naturally occurring amino acid.

Amino acids may be referred to herein by either their commonly known three letter symbols or by the one-letter symbols recommended by the IUPAC-IUB Biochemical Nomenclature Commission. Nucleotides, likewise, may be referred to by their commonly accepted single-letter codes.

"Conservatively modified variants" applies to both amino acid and nucleic acid sequences. With respect to particular nucleic acid sequences, conservatively modified variants refers to those nucleic acids which encode identical or essentially identical amino acid sequences, or where the nucleic acid does not encode an amino acid sequence, to essentially identical or associated, e.g., naturally contiguous, sequences. Because of the degeneracy of the genetic code, a large number of functionally identical nucleic acids encode most proteins. For instance, the codons GCA, GCC, GCG, and GCU encode the amino acid alanine. Thus, at every position where an alanine is specified by a codon, the codon can be altered to another of the corresponding codons described without altering the encoded polypeptide. Such nucleic acid variations are "silent variations," which are one species of conservatively modified variations. Every nucleic acid sequence herein which encodes a polypeptide also describes silent variations of the nucleic acid. One of skill will recognize that in certain contexts each codon in a nucleic acid (except AUG, which is ordinarily the only codon for methionine, and TGG, which is ordinarily the only codon for tryptophan) can be modified to yield a functionally identical molecule. Accordingly, often silent variations of

a nucleic acid which encodes a polypeptide is implicit in a described sequence with respect to the expression product, but not with respect to actual probe sequences.

As to amino acid sequences, one of skill will recognize that individual substitutions, deletions or additions to a nucleic acid, peptide, polypeptide, or protein sequence which alters, adds or deletes a single amino acid or a small percentage of amino acids in the encoded sequence is a "conservatively modified variant" where the alteration results in the substitution of an amino acid with a chemically similar amino acid. Conservative substitutions providing functionally similar amino acids are well known in the art. Such conservatively modified variants are in addition to and do not exclude polymorphic variants, interspecies homologs, and alleles of the invention, typically conservative substitutions for one another: 1) Alanine (A), Glycine (G); 2) Aspartic acid (D), Glutamic acid (E); 3) Asparagine (N), Glutamine (Q); 4) Arginine (R), Lysine (K); 5) Isoleucine (I), Leucine (L), Methionine (M), Valine (V); 6) Phenylalanine (F), Tyrosine (Y), Tryptophan (W); 7) Serine (S), Threonine (T); and 8) Cysteine (C), Methionine (M) (see, e.g., Creighton (1984) Proteins Freeman).

Macromolecular structures such as polypeptide structures can be described in terms of various levels of organization. For a general discussion of this organization, see, e.g., Alberts, et al. (2001) Molecular Biology of the Cell (4th ed.) and Cantor and Schimmel (1980) Biophysical Chemistry Part I: The Conformation of Biological Macromolecules Freeman. "Primary structure" refers to the amino acid sequence of a particular peptide. "Secondary structure" refers to locally ordered, three dimensional structures within a polypeptide. These structures are commonly known as domains. Domains are portions of a polypeptide that often form a compact unit of the polypeptide and are typically 25 to approximately 500 amino acids long. Typical domains are made up of sections of lesser organization such as stretches of  $\beta$ -sheet and  $\alpha$ -helices. "Tertiary structure" refers to the complete three dimensional structure of a polypeptide monomer. "Quaternary structure" refers to the three dimensional structure formed, usually by the noncovalent association of independent tertiary units. Anisotropic terms are also known as energy terms.

"Nucleic acid" or "oligonucleotide" or "polynucleotide" or grammatical equivalents used herein means at least two nucleotides covalently linked together. Oligonucleotides are typically from about 5, 6, 7, 8, 9, 10, 12, 15, 25, 30, 40, 50 or more nucleotides in length, up to about 100 nucleotides in length. Nucleic acids and polynucleotides are a polymers of virtually any length, including longer lengths, e.g., 200, 300, 500, 1000, 2000, 3000, 5000,

- 7000, 10,000, etc. A nucleic acid of the present invention will generally contain phosphodiester bonds, although in some cases, nucleic acid analogs are included that may have alternate backbones, comprising, e.g., phosphoramidate, phosphorothioate, phosphorodithioate, or O-methylphosphoroamidite linkages (see Eckstein (1992) Oligonucleotides and Analogues: A Practical Approach, Oxford University Press); and peptide nucleic acid backbones and linkages. Other analog nucleic acids include those with positive backbones; non-ionic backbones, and non-ribose backbones, including those described in U.S. Patent Nos. 5,235,033 and 5,034,506, and Chapters 6 and 7, ASC Symposium Series 580, Sanghvi and Cook (eds. 1994) Carbohydrate Modifications in Antisense Research ACS Symposium Series 580. Nucleic acids containing one or more carbocyclic sugars are also included within one definition of nucleic acids. Modifications of the ribose-phosphate backbone may be done for a variety of reasons, e.g., to increase the stability and half-life of such molecules in physiological environments or as probes on a biochip. Mixtures of naturally occurring nucleic acids and analogs can be made; alternatively, mixtures of different nucleic acid analogs, and mixtures of naturally occurring nucleic acids and analogs may be made.

- A variety of references disclose such nucleic acid analogs, including, for example, phosphoramidate (Beaucage, et al. (1993) Tetrahedron 49(10):1925-1963 and references therein; Letsinger (1970) J. Org. Chem. 35:3800-3803; Sprinzl, et al. (1977) Eur. J. Biochem. 81:579-589; Letsinger, et al. (1986) Nucl. Acids Res. 14:3487-499; Sawai, et al (1984) Chem. Lett. 805; Letsinger, et al. (1988) J. Am. Chem. Soc. 110:4470-4471; and Pauwels, et al. (1986) Chemica Scripta 26:141-149), phosphorothioate (Mag, et al. (1991) Nucleic Acids Res. 19:1437-441; and U.S. Patent No. 5,644,048), phosphorodithioate (Briu, et al. (1989) J. Am. Chem. Soc. 111:2321-xxx, O-methylphosphoroamidite linkages (see Eckstein (1992) Oligonucleotides and Analogues: A Practical Approach Oxford University Press), and peptide nucleic acid backbones and linkages (see Egholm (1992) J. Am. Chem. Soc. 114:1895-1897; Meier, et al. (1992) Chem. Int. Ed. Engl. 31:1008-1010; Nielsen (1993) Nature 365:566-568; Carlsson, et al. (1996) Nature 380:207, each of which is incorporated by reference). Other analog nucleic acids include those with positive backbones (Denpcey, et al. (1995) Proc. Natl. Acad. Sci. USA 92:6097-101; non-ionic backbones (U.S. Patent Nos. 5,386,023, 5,637,684, 5,602,240, 5,216,141 and 4,469,863; Kiedrowski, et al. (1991) Angew. Chem. Intl. Ed. English 30:423-426; Letsinger, et al. (1988) J. Am. Chem. Soc. 110:4470;

Letsinger, et al. (1994) Nucleoside and Nucleotide 13:1597-xxx; Chapters 2 and 3 in Sanghvi and Cook (eds. 1994) Carbohydrate Modifications in Antisense Research ACS Symposium Series 580; Mesmaeker, et al. (1994) Bioorganic and Medicinal Chem. Lett. 4:395-xxx; Jeffs, et al. (1994) J. Biomolecular NMR 34:17; Horn (1996) Tetrahedron Lett. 37:743-xxx) and non-ribose backbones, including those described in U.S. Patent Nos. 5,235,033 and 5,034,506, and Chapters 6 and 7 in Sanghvi and Cook (eds. 1994) Carbohydrate Modifications in Antisense Research ACS Symposium Series 580. Nucleic acids containing one or more carbocyclic sugars are also included within one definition of nucleic acids (see Jenkins, et al. (1995) Chem. Soc. Rev. xx:169-176). Several nucleic acid analogs are described in Rawls (p. 35, June 2, 1997) C&E News. Each of these references is hereby expressly incorporated by reference.

Particularly preferred are peptide nucleic acids (PNA) which includes peptide nucleic acid analogs. These backbones are substantially non-ionic under neutral conditions, in contrast to the highly charged phosphodiester backbone of naturally occurring nucleic acids. This results in two advantages. First, the PNA backbone exhibits improved hybridization kinetics. PNAs have larger changes in the melting temperature ( $T_m$ ) for mismatched versus perfectly matched base pairs. DNA and RNA typically exhibit a 2-4° C drop in  $T_m$  for an internal mismatch. With the non-ionic PNA backbone, the drop is closer to 7-9° C. Similarly, due to their non-ionic nature, hybridization of the bases attached to these backbones is relatively insensitive to salt concentration. In addition, PNAs are not degraded by cellular enzymes, and thus can be more stable.

The nucleic acids may be single stranded or double stranded, as specified, or contain portions of both double stranded or single stranded sequence. As will be appreciated by those in the art, the depiction of a single strand also defines the sequence of the complementary strand; thus the sequences described herein also provide the complement of the sequence. The nucleic acid may be DNA, both genomic and cDNA, RNA or a hybrid, where the nucleic acid may contain combinations of deoxyribo- and ribo-nucleotides, and combinations of bases, including uracil, adenine, thymine, cytosine, guanine, inosine, xanthine hypoxanthine, isocytosine, isoguanine, etc. "Transcript" typically refers to a naturally occurring RNA, e.g., a pre-mRNA, hnRNA, or mRNA. As used herein, the term "nucleoside" includes nucleotides and nucleoside and nucleotide analogs, and modified nucleosides such as amino modified nucleosides. In addition, "nucleoside" includes non-naturally occurring analog structures.

Thus, e.g., the individual units of a peptide nucleic acid, each containing a base, are referred to herein as a nucleoside.

A "label" or a "detectable moiety" is a composition detectable by spectroscopic, photochemical, biochemical, immunochemical, chemical, or other physical means. For example, useful labels include  $^{32}\text{P}$ , fluorescent dyes, electron-dense reagents, enzymes (e.g., as commonly used in an ELISA), biotin, digoxigenin, or haptens and proteins or other entities which can be made detectable, e.g., by incorporating a radiolabel into the peptide or used to detect antibodies specifically reactive with the peptide. The labels may be incorporated into the prostate cancer nucleic acids, proteins, and antibodies at virtually any position. Many methods for conjugating the antibody to the label may be employed, including those methods described by Hunter, et al. (1962) Nature, 144:945; David, et al. (1974) Biochemistry 13:1014-1021; Pain, et al. (1981) J. Immunol. Meth. 40:219-230; and Nygren (1982) J. Histochem. and Cytochem. 30:407-412.

An "effector" or "effector moiety" or "effector component" is a molecule that is bound (or linked, or conjugated), either covalently, through a linker or a chemical bond, or noncovalently, through ionic, van der Waals, electrostatic, or hydrogen bonds, to an antibody. The "effector" can be a variety of molecules including, e.g., detection moieties including radioactive compounds, fluorescent compounds, an enzyme or substrate, tags such as epitope tags, a toxin; activatable moieties, a chemotherapeutic agent; a lipase; an antibiotic; or a radioisotope emitting "hard" e.g., beta radiation.

A "labeled nucleic acid probe or oligonucleotide" is one that is bound, either covalently, through a linker or a chemical bond, or noncovalently, through ionic, van der Waals, electrostatic, or hydrogen bonds to a label such that the presence of the probe may be detected by detecting the presence of the label bound to the probe. Alternatively, method using high affinity interactions may achieve the same results where one of a pair of binding partners binds to the other, e.g., biotin, streptavidin.

As used herein a "nucleic acid probe or oligonucleotide" is defined as a nucleic acid capable of binding to a target nucleic acid of complementary sequence through one or more types of chemical bonds, usually through complementary base pairing, usually through hydrogen bond formation. As used herein, a probe may include natural (i.e., A, G, C, or T) or modified bases (7-deazaguanosine, inosine, etc.). In addition, the bases in a probe may be joined by a linkage other than a phosphodiester bond, so long as it does not functionally



interfere with hybridization. Thus, e.g., probes may be peptide nucleic acids in which the constituent bases are joined by peptide bonds rather than phosphodiester linkages. It will be understood by one of skill in the art that probes may bind target sequences lacking complete complementarity with the probe sequence depending upon the stringency of the hybridization conditions. The probes are preferably directly labeled as with isotopes, chromophores, lumiphores, chromogens, or indirectly labeled such as with biotin to which a streptavidin complex may later bind. By assaying for the presence or absence of the probe, one can detect the presence or absence of the select sequence or subsequence. Diagnosis or prognosis may be based at the genomic level, or at the level of RNA or protein expression.

The term "recombinant" when used with reference, e.g., to a cell, or nucleic acid, protein, or vector, indicates that the cell, nucleic acid, protein or vector, has been modified by the introduction of a heterologous nucleic acid or protein or the alteration of a native nucleic acid or protein, or that the cell is derived from a cell so modified. Thus, e.g., recombinant cells express genes that are not found within the native (non-recombinant) form of the cell or express native genes that are otherwise abnormally expressed, under expressed or not expressed at all. By the term "recombinant nucleic acid" herein is meant nucleic acid, originally formed in vitro, in general, by the manipulation of nucleic acid, e.g., using polymerases and endonucleases, in a form not normally found in nature. In this manner, operably linkage of different sequences is achieved. Thus an isolated nucleic acid, in a linear form, or an expression vector formed in vitro by ligating DNA molecules that are not normally joined, are both considered recombinant for the purposes of this invention. It is understood that once a recombinant nucleic acid is made and reintroduced into a host cell or organism, it will replicate non-recombinantly, i.e., using the in vivo cellular machinery of the host cell rather than in vitro manipulations; however, such nucleic acids, once produced recombinantly, although subsequently replicated non-recombinantly, are still considered recombinant for the purposes of the invention. Similarly, a "recombinant protein" is a protein made using recombinant techniques, i.e., through the expression of a recombinant nucleic acid as depicted above.

The term "heterologous" when used with reference to portions of a nucleic acid indicates that the nucleic acid comprises two or more subsequences that are not normally found in the same relationship to each other in nature. For instance, the nucleic acid is typically recombinantly produced, having two or more sequences, e.g., from unrelated genes

arranged to make a new functional nucleic acid, e.g., a promoter from one source and a coding region from another source. Similarly, a heterologous protein will often refer to two or more subsequences that are not found in the same relationship to each other in nature (e.g., a fusion protein).

5 A "promoter" is defined as an array of nucleic acid control sequences that direct transcription of a nucleic acid. As used herein, a promoter includes necessary nucleic acid sequences near the start site of transcription, such as, in the case of a polymerase II type promoter, a TATA element. A promoter also optionally includes distal enhancer or repressor elements, which can be located as much as several thousand base pairs from the start site of  
10 transcription. A "constitutive" promoter is a promoter that is active under most environmental and developmental conditions. An "inducible" promoter is a promoter that is active under environmental or developmental regulation. The term "operably linked" refers to a functional linkage between a nucleic acid expression control sequence (such as a promoter, or array of transcription factor binding sites) and a second nucleic acid sequence,  
15 wherein the expression control sequence directs transcription of the nucleic acid corresponding to the second sequence.

An "expression vector" is a nucleic acid construct, generated recombinantly or synthetically, with a series of specified nucleic acid elements that permit transcription of a particular nucleic acid in a host cell. The expression vector can be part of a plasmid, virus, or  
20 nucleic acid fragment. Typically, the expression vector includes a nucleic acid to be transcribed operably linked to a promoter.

The phrase "selectively (or specifically) hybridizes to" refers to the binding, duplexing, or hybridizing of a molecule only to a particular nucleotide sequence under stringent hybridization conditions when that sequence is present in a complex mixture (e.g.,  
25 total cellular or library DNA or RNA).

The phrase "stringent hybridization conditions" refers to conditions under which a probe will hybridize to its target subsequence, typically in a complex mixture of nucleic acids, but to no other sequences. Stringent conditions are sequence-dependent and will be different in different circumstances. Longer sequences hybridize specifically at higher  
30 temperatures. An extensive guide to the hybridization of nucleic acids is found "Overview of principles of hybridization and the strategy of nucleic acid assays" in Tijssen (1993) Hybridization with Nucleic Probes (Techniques in Biochemistry and Molecular Biology vol.

24) Elsevier. Generally, stringent conditions are selected to be about 5-10° C lower than the thermal melting point ( $T_m$ ) for the specific sequence at a defined ionic strength pH. The  $T_m$  is the temperature (under defined ionic strength, pH, and nucleic concentration) at which 50% of the probes complementary to the target hybridize to the target sequence at equilibrium (as the target sequences are present in excess, at  $T_m$ , 50% of the probes are occupied at equilibrium). Stringent conditions will be those in which the salt concentration is less than about 1.0 M sodium ion, typically about 0.01 to 1.0 M sodium ion concentration (or other salts) at pH 7.0 to 8.3 and the temperature is at least about 30° C for short probes (e.g., 10 to 50 nucleotides) and at least about 60° C for long probes (e.g., greater than 50 nucleotides). Stringent conditions may also be achieved with the addition of destabilizing agents such as formamide. For selective or specific hybridization, a positive signal is at least two times background, preferably 10 times background hybridization. Exemplary stringent hybridization conditions can be as following: 50% formamide, 5x SSC, and 1% SDS, incubating at 42° C, or, 5x SSC, 1% SDS, incubating at 65° C, with wash in 0.2x SSC, and 0.1% SDS at 65° C. For PCR, a temperature of about 36° C is typical for low stringency amplification, although annealing temperatures may vary between about 32° C and 48° C depending on primer length. For high stringency PCR amplification, a temperature of about 62° C is typical, although high stringency annealing temperatures can range from about 50-65° C, depending on the primer length and specificity. Typical cycle conditions for both high and low stringency amplifications include a denaturation phase of 90-95° C for 30-120 sec, an annealing phase lasting 30-120 sec, and an extension phase of about 72° C for 1-2 min. Protocols and guidelines for low and high stringency amplification reactions are provided, e.g., in Innis, et al. (1990) PCR Protocols: A Guide to Methods and Applications Academic Press, N.Y.

Nucleic acids that do not hybridize to each other under stringent conditions are still substantially identical if the polypeptides which they encode are substantially identical. This occurs, e.g., when a copy of a nucleic acid is created using the maximum codon degeneracy permitted by the genetic code. In such cases, the nucleic acids typically hybridize under moderately stringent hybridization conditions. Exemplary "moderately stringent hybridization conditions" include a hybridization in a buffer of 40% formamide, 1 M NaCl, 1% SDS at 37° C, and a wash in 1X SSC at 45° C. A positive hybridization is at least twice

background. Those of ordinary skill will readily recognize that alternative hybridization and wash conditions can be utilized to provide conditions of similar stringency. Additional guidelines for determining hybridization parameters are provided in numerous references, e.g., Ausubel, et al. (eds. 1991 and supplements) Current Protocols in Molecular Biology

5       The phrase "functional effects" in the context of assays for testing compounds that modulate activity of a prostate cancer protein includes the determination of a parameter that is indirectly or directly under the influence of the prostate cancer protein or nucleic acid, e.g., a functional, physical, or chemical effect, such as the ability to decrease prostate proliferation (malignant or non-malignant). It includes ligand binding activity; cell growth on soft agar;  
10   anchorage dependence; contact inhibition and density limitation of growth; cellular proliferation; cellular transformation; growth factor or serum dependence; tumor specific marker levels; invasiveness into Matrigel; tumor growth and metastasis in vivo; mRNA and protein expression in cells undergoing metastasis, and other characteristics of prostate cancer cells. "Functional effects" include in vitro, in vivo, and ex vivo activities.

15       By "determining the functional effect" is meant assaying for a compound that increases or decreases a parameter that is indirectly or directly under the influence of a prostate cancer protein sequence, e.g., functional, enzymatic, physical and chemical effects. Such functional effects can be measured by means known to those skilled in the art, e.g., changes in spectroscopic characteristics (e.g., fluorescence, absorbance, refractive index),  
20   hydrodynamic (e.g., shape), chromatographic, or solubility properties for the protein, measuring inducible markers or transcriptional activation of the prostate cancer protein; measuring binding activity or binding assays, e.g., binding to antibodies or other ligands, and measuring cellular proliferation. Determination of the functional effect of a compound on prostate cancer can also be performed using prostate cancer assays known to those of skill in  
25   the art such as an in vitro assays, e.g., cell growth on soft agar; anchorage dependence; contact inhibition and density limitation of growth; cellular proliferation; cellular transformation; growth factor or serum dependence; tumor specific marker levels; invasiveness into Matrigel; tumor growth and metastasis in vivo; mRNA and protein expression in cells undergoing metastasis, and other characteristics of prostate cancer cells.  
30   The functional effects can be evaluated by many means known to those skilled in the art, e.g., microscopy for quantitative or qualitative measures of alterations in morphological features, measurement of changes in RNA or protein levels for prostate cancer-associated sequences,

measurement of RNA stability, identification of downstream or reporter gene expression (CAT, luciferase,  $\beta$ -gal, GFP, and the like), e.g., via chemiluminescence, fluorescence, colorimetric reactions, antibody binding, inducible markers, and ligand binding assays.

“Inhibitors”, “activators”, and “modulators” of prostate cancer polynucleotide and polypeptide sequences are used to refer to activating, inhibitory, or modulating molecules or compounds identified using in vitro and in vivo assays of prostate cancer polynucleotide and polypeptide sequences. Inhibitors are compounds that, e.g., bind to, partially or totally block activity, decrease, prevent, delay activation, inactivate, desensitize, or down regulate the activity or expression of prostate cancer proteins, e.g., antagonists. Antisense nucleic acids may seem to inhibit expression and subsequent function of the protein. “Activators” are compounds that increase, open, activate, facilitate, enhance activation, sensitize, agonize, or up regulate prostate cancer protein activity. Inhibitors, activators, or modulators also include genetically modified versions of prostate cancer proteins, e.g., versions with altered activity, as well as naturally occurring and synthetic ligands, antagonists, agonists, antibodies, small chemical molecules and the like. Such assays for inhibitors and activators include, e.g., expressing the prostate cancer protein in vitro, in cells, or cell membranes, applying putative modulator compounds, and then determining the functional effects on activity, as described above. Activators and inhibitors of prostate cancer can also be identified by incubating prostate cancer cells with the test compound and determining increases or decreases in the expression of 1 or more prostate cancer proteins, e.g., 1, 2, 3, 4, 5, 10, 15, 20, 25, 30, 40, 50 or more prostate cancer proteins, such as prostate cancer proteins encoded by the sequences set out in Tables 1A-4.

Samples or assays comprising prostate cancer proteins that are treated with a potential activator, inhibitor, or modulator are compared to control samples without the inhibitor, activator, or modulator to examine the extent of inhibition. Control samples (untreated with inhibitors) are assigned a relative protein activity value of 100%. Inhibition of a polypeptide is achieved when the activity value relative to the control is about 80%, preferably 50%, more preferably 25-0%. Activation of a prostate cancer polypeptide is achieved when the activity value relative to the control (untreated with activators) is 110%, more preferably 150%, more preferably 200-500% (i.e., two to five fold higher relative to the control), more preferably 1000-3000% higher.

The phrase “changes in cell growth” refers to a change in cell growth and proliferation characteristics in vitro or in vivo, such as cell viability, formation of foci, anchorage independence, semi-solid or soft agar growth, changes in contact inhibition and density limitation of growth, loss of growth factor or serum requirements, changes in cell morphology, gaining or losing immortalization, gaining or losing tumor specific markers, ability to form or suppress tumors when injected into suitable animal hosts, and/or immortalization of the cell. See, e.g., pp. 231-241 in Freshney (1994) Culture of Animal Cells: A Manual of Basic Technique (3d ed.) Wiley-Liss.

“Tumor cell” refers to precancerous, cancerous, and/or normal cells in a tumor.

“Cancer cells,” “transformed” cells, or “transformation” in tissue culture, refers to spontaneous or induced phenotypic changes that do not necessarily involve the uptake of new genetic material. Although transformation can arise from infection with a transforming virus and incorporation of new genomic DNA, or uptake of exogenous DNA, it can also arise spontaneously or following exposure to a carcinogen, thereby mutating an endogenous gene. Transformation is associated with phenotypic changes, such as immortalization of cells, aberrant growth control, nonmorphological changes, and/or malignancy. See, Freshney (2001) Culture of Animal Cells: A Manual of Basic Technique (4th ed.) Wiley-Liss.

“Antibody” refers to a polypeptide comprising a framework region from an immunoglobulin gene or fragments thereof that specifically binds and recognizes an antigen.

The recognized immunoglobulin genes include the kappa, lambda, alpha, gamma, delta, epsilon, and mu constant region genes, as well as the myriad immunoglobulin variable region genes. Light chains are classified as either kappa or lambda. Heavy chains are classified as gamma, mu, alpha, delta, or epsilon, which in turn define the immunoglobulin classes, IgG, IgM, IgA, IgD, and IgE, respectively. Typically, the antigen-binding region of an antibody or its functional equivalent will be most critical in specificity and affinity of binding. See Paul (ed. 1999) Fundamental Immunology (4th ed.) Raven.

An exemplary immunoglobulin (antibody) structural unit comprises a tetramer. Each tetramer is composed of two identical pairs of polypeptide chains, each pair having one “light” (about 25 kD) and one “heavy” chain (about 50-70 kD). The N-terminus of each chain defines a variable region of about 100 to 110 or more amino acids primarily responsible for antigen recognition. The terms variable light chain ( $V_L$ ) and variable heavy chain ( $V_H$ ) refer to these light and heavy chains respectively.

Antibodies exist, e.g., as intact immunoglobulins or as a number of well-characterized fragments produced by digestion with various peptidases. Thus, e.g., pepsin digests an antibody below the disulfide linkages in the hinge region to produce  $F(ab)'_2$ , a dimer of Fab which itself is a light chain joined to  $V_H-C_H1$  by a disulfide bond. The  $F(ab)'_2$  may be reduced under mild conditions to break the disulfide linkage in the hinge region, thereby converting the  $F(ab)'_2$  dimer into an Fab' monomer. The Fab' monomer is essentially Fab with part of the hinge region (see Paul (ed. 1993) Fundamental Immunology (3d ed.) Raven. While various antibody fragments are defined in terms of the digestion of an intact antibody, one of skill will appreciate that such fragments may be synthesized de novo either chemically or by using recombinant DNA methodology. Thus, the term antibody, as used herein, also includes antibody fragments either produced by the modification of whole antibodies, or those synthesized de novo using recombinant DNA methodologies (e.g., single chain Fv) or those identified using phage display libraries (see, e.g., McCafferty, et al. (1990) Nature 348:552-554.

For preparation of antibodies, e.g., recombinant, monoclonal, or polyclonal antibodies, many technique known in the art can be used (see, e.g., Kohler and Milstein (1975) Nature 256:495-497; Kozbor, et al. (1983) Immunology Today 4:72; pp. 77-96 in Cole, et al. (1985) Monoclonal Antibodies and Cancer Therapy Liss; Coligan (1991) Current Protocols in Immunology Lippincott; Harlow and Lane (1988) Antibodies: A Laboratory Manual CSH Press; and Goding (1986) Monoclonal Antibodies: Principles and Practice (2d ed.) Academic Press. Techniques for the production of single chain antibodies (U.S. Patent 4,946,778) can be adapted to produce antibodies to polypeptides of this invention. Also, transgenic mice, or other organisms such as other mammals, may be used to express humanized antibodies. Alternatively, phage display technology can be used to identify antibodies and heteromeric Fab fragments that specifically bind to selected antigens (see, e.g., McCafferty, et al. (1990) Nature 348:552-554; Marks, et al. (1992) Biotechnology 10:779-783).

A "chimeric antibody" is an antibody molecule in which (a) the constant region, or a portion thereof, is altered, replaced or exchanged so that the antigen binding site (variable region) is linked to a constant region of a different or altered class, effector function and/or species, or an entirely different molecule which confers new properties to the chimeric antibody, e.g., an enzyme, toxin, hormone, growth factor, drug, etc.; or (b) the variable

region, or a portion thereof, is altered, replaced or exchanged with a variable region having a different or altered antigen specificity.

#### Identification of prostate cancer-associated sequences

5 In one aspect, the expression levels of genes are determined in different patient samples for which diagnosis information is desired, to provide expression profiles. An expression profile of a particular sample is essentially a "fingerprint" of the state of the sample; while two states may have a particular gene similarly expressed, the evaluation of a number of genes simultaneously allows the generation of a gene expression profile that is  
10 characteristic of the state of the cell. That is, normal tissue (e.g., normal prostate or other tissue) may be distinguished from pathological prostate cells, e.g., cancerous or metastatic cancerous tissue of the prostate, or prostate cancer tissue or metastatic prostate cancerous tissue can be compared with tissue samples of prostate and other tissues from surviving cancer patients. By comparing expression profiles of tissue in known different prostate  
15 cancer states, information regarding which genes are important (including both up- and down-regulation of genes) in each of these states is obtained.

The identification of sequences that are differentially expressed in prostate cancer versus non-prostate cancer tissue allows the use of this information in a number of ways. For example, a particular treatment regime may be evaluated: does a chemotherapeutic drug act  
20 to down-regulate prostate cancer or other proliferative disorders, and thus tumor growth or recurrence, in a particular patient. Alternatively, a treatment step may induce other markers which may be used as targets to destroy tumor cells. Similarly, diagnosis and treatment outcomes may be done or confirmed by comparing patient samples with the known expression profiles. Malignant disease may be compared to non-malignant conditions.  
25 Metastatic tissue can also be analyzed to determine the stage of prostate cancer in the tissue, or origin of primary tumor, e.g., metastasis from a remote primary site. Furthermore, these gene expression profiles (or individual genes) allow screening of drug candidates with an eye to mimicking or altering a particular expression profile; e.g., screening can be done for drugs that suppress the prostate cancer expression profile. This may be done by making biochips  
30 comprising sets of the important prostate cancer genes, which can then be used in these screens. These methods can also be done on the protein basis; that is, protein expression levels of the prostate cancer proteins can be evaluated for diagnostic purposes or to screen



candidate agents. In addition, the prostate cancer nucleic acid sequences can be administered for gene therapy purposes, including the administration of antisense nucleic acids, or the prostate cancer proteins (including antibodies and other modulators thereof) administered as therapeutic drugs.

5 Thus the present invention provides nucleic acid and protein sequences that are differentially expressed in prostate cancer relative to normal tissues and/or non-malignant disease, or in different types of related diseases, herein termed "prostate cancer sequences." As outlined below, prostate cancer sequences include those that are up-regulated (i.e., expressed at a higher level) in prostate cancer, as well as those that are down-regulated (i.e.,  
10 expressed at a lower level). In a preferred embodiment, the prostate cancer sequences are from humans; however, as will be appreciated by those in the art, prostate cancer sequences from other organisms may be useful in animal models of disease and drug evaluation; thus, other prostate cancer sequences are provided, from vertebrates, including mammals, including rodents (rats, mice, hamsters, guinea pigs, etc.), primates, farm animals (including  
15 sheep, goats, pigs, cows, horses, etc.) and pets, e.g., (dogs, cats, etc.). Prostate cancer sequences from other organisms may be obtained using the techniques outlined below.

Prostate cancer sequences can include both nucleic acid and amino acid sequences. As will be appreciated by those in the art and is more fully outlined below, prostate cancer nucleic acid sequences are useful in a variety of applications, including diagnostic  
20 applications, which will detect naturally occurring nucleic acids, as well as screening applications; e.g., biochips comprising nucleic acid probes or PCR microtiter plates with selected probes to the prostate cancer sequences can be generated.

A prostate cancer sequence can be initially identified by substantial nucleic acid and/or amino acid sequence homology to the prostate cancer sequences outlined herein. Such  
25 homology can be based upon the overall nucleic acid or amino acid sequence, and is generally determined as outlined below, using either homology programs or hybridization conditions.

For identifying prostate cancer-associated sequences, the prostate cancer screen typically includes comparing genes identified in different tissues, e.g., normal and cancerous  
30 tissues, or tumor tissue samples from patients who have metastatic disease vs. non metastatic tissue. Other suitable tissue comparisons include comparing prostate cancer samples with metastatic cancer samples from other cancers, such as lung, breast, gastrointestinal cancers,

ovarian, etc. Samples of different stages of prostate cancer, e.g., survivor tissue, drug resistant states, and tissue undergoing metastasis, are applied to biochips comprising nucleic acid probes. The samples are first microdissected, if applicable, and treated as is known in the art for the preparation of mRNA. Suitable biochips are commercially available, e.g., from  
5 Affymetrix. Gene expression profiles are generated and the data analyzed.

In one embodiment, the genes showing changes in expression as between normal and disease states are compared to genes expressed in other normal tissues, preferably normal prostate, but also including, and not limited to lung, heart, brain, liver, breast, kidney, muscle, colon, small intestine, large intestine, spleen, bone, and placenta. In a preferred embodiment,  
10 those genes identified during the prostate cancer screen that are expressed in a significant amount in other tissues are removed from the profile, although in some embodiments, this is not necessary. That is, when screening for drugs, it is usually preferable that the target be disease specific, to minimize possible side effects on other organs were there expression.

In a preferred embodiment, prostate cancer sequences are those that are up-regulated  
15 in prostate cancer or related conditions; that is, the expression of these genes is higher in the prostate cancer tissue as compared to non-cancerous tissue. "Up-regulation" as used herein often means at least about a two-fold change, preferably at least about a three fold change, with at least about five-fold or higher being preferred. Another embodiment is directed to sequences up-regulated in non-malignant conditions relative to normal.

Unigene cluster identification numbers and accession numbers herein are for the  
20 GenBank sequence database and the sequences of the accession numbers are hereby expressly incorporated by reference. GenBank is known in the art, see, e.g., Benson, et al. (1998) Nucleic Acids Research 26:1-7 and <http://www.ncbi.nlm.nih.gov/>. Sequences are also available in other databases, e.g., European Molecular Biology Laboratory (EMBL) and  
25 DNA Database of Japan (DDBJ). U.S. Patent Application N. 09/687,576 and 09/976,858 (-001-3) further disclose related sequences, compositions, and methods of diagnosis and treatment of prostate cancer and related conditions and are hereby expressly incorporated by reference.

In another preferred embodiment, prostate cancer sequences are those that are down-  
30 regulated in the prostate cancer; that is, the expression of these genes is lower in prostate cancer tissue as compared to non-cancerous tissue. "Down-regulation" as used herein often

means at least about a two-fold change, preferably at least about a three fold change, with at least about five-fold or higher being preferred.

#### Informatics

5       The ability to identify genes that are over or under expressed in prostate cancer can additionally provide high-resolution, high-sensitivity datasets which can be used in the areas of diagnostics, therapeutics, drug development, pharmacogenetics, protein structure, biosensor development, and other related areas. For example, the expression profiles can be used in diagnostic or prognostic evaluation of patients with prostate cancer. Or as another  
10       example, subcellular toxicological information can be generated to better direct drug structure and activity correlation (see Anderson, Pharmaceutical Proteomics: Targets, Mechanism, and Function, paper presented at the IBC Proteomics conference, Coronado, CA (June 11-12, 1998)). Subcellular toxicological information can also be utilized in a biological sensor device to predict the likely toxicological effect of chemical exposures and likely tolerable  
15       exposure thresholds (see U.S. Patent No. 5,811,231). Similar advantages accrue from datasets relevant to other biomolecules and bioactive agents (e.g., nucleic acids, saccharides, lipids, drugs, and the like).

      Thus, in another embodiment, the present invention provides a database that includes at least one set of assay data. The data contained in the database is acquired, e.g., using array  
20       analysis either singly or in a library format. The database can be in a form in which data can be maintained and transmitted, but is preferably an electronic database. The electronic database of the invention can be maintained on an electronic device allowing for the storage of and access to the database, such as a personal computer, but is preferably distributed on a wide area network, such as the World Wide Web.

25       The focus of the present section on databases that include peptide sequence data is for clarity of illustration only. It will be apparent to those of skill in the art that similar databases can be assembled for assay data acquired using an assay of the invention.

      The compositions and methods for identifying and/or quantitating the relative and/or absolute abundance of a variety of molecular and macromolecular species from a biological  
30       sample undergoing prostate cancer, i.e., the identification of prostate cancer-associated sequences described herein, provide an abundance of information, which can be correlated with pathological conditions, predisposition to disease, drug testing, therapeutic monitoring,

gene-disease causal linkages, identification of correlates of immunity and physiological status, among others. Although the data generated from the assays of the invention is suited for manual review and analysis, in a preferred embodiment, prior data processing using high-speed computers is utilized.

5 An array of methods for indexing and retrieving biomolecular information is known in the art. For example, U.S. Patents 6,023,659 and 5,966,712 disclose a relational database system for storing biomolecular sequence information in a manner that allows sequences to be catalogued and searched according to one or more protein function hierarchies. U.S. Patent 5,953,727 discloses a relational database having sequence records containing  
10 information in a format that allows a collection of partial-length DNA sequences to be catalogued and searched according to association with one or more sequencing projects for obtaining full-length sequences from the collection of partial length sequences. U.S. Patent 5,706,498 discloses a gene database retrieval system for making a retrieval of a gene sequence similar to a sequence data item in a gene database based on the degree of similarity  
15 between a key sequence and a target sequence. U.S. Patent 5,538,897 discloses a method using mass spectroscopy fragmentation patterns of peptides to identify amino acid sequences in computer databases by comparison of predicted mass spectra with experimentally-derived mass spectra using a closeness-of-fit measure. U.S. Patent 5,926,818 discloses a multi-dimensional database comprising a functionality for multi-dimensional data analysis  
20 described as on-line analytical processing (OLAP), which entails the consolidation of projected and actual data according to more than one consolidation path or dimension. U.S. Patent 5,295,261 reports a hybrid database structure in which the fields of each database record are divided into two classes, navigational and informational data, with navigational fields stored in a hierarchical topological map which can be viewed as a tree structure or as  
25 the merger of two or more such tree structures.

See also Mount, et al. (2001) Bioinformatics CSH Press; Durbin, et al. (eds. 1999) Biological Sequence Analysis: Probabilistic Models of Proteins and Nucleic Acids Cambridge Univ. Press; Baxeavanis and Ouellette (eds., 1998) Bioinformatics: A Practical Guide to the Analysis of Genes and Proteins Wiley-Liss; Rashidi and Buehler (1999)  
30 Bioinformatics: Basic Applications in Biological Science and Medicine CRC Press; Setubal, et al. (eds. 1997) Introduction to Computational Molecular Biology Brooks/Cole; Misener and Krawetz (eds. 2000) Bioinformatics: Methods and Protocols Human Press; Higgins and

- Taylor (eds. 2000) Bioinformatics: Sequence, Structure, and Databanks: A Practical Approach Oxford Univ. Press; Brown (2001) Bioinformatics: A Biologist's Guide to Biocomputing and the Internet Eaton Pub; Han and Kamber (2000) Data Mining: Concepts and Techniques Kaufmann Pub.; and Waterman (1995) Introduction to Computational Biology: Maps, Sequences, and Genomes Chap and Hall.

The present invention provides a computer database comprising a computer and software for storing in computer-retrievable form assay data records cross-tabulated, e.g., with data specifying the source of the target-containing sample from which each sequence specificity record was obtained.

- In an exemplary embodiment, at least one of the sources of target-containing sample is from a control tissue sample known to be free of pathological disorders. In a variation, at least one of the sources is a known pathological tissue specimen, e.g., a neoplastic lesion or another tissue specimen to be analyzed for prostate cancer. In another variation, the assay records cross-tabulate one or more of the following parameters for each target species in a sample: (1) a unique identification code, which can include, e.g., a target molecular structure and/or characteristic separation coordinate (e.g., electrophoretic coordinates); (2) sample source; and (3) absolute and/or relative quantity of the target species present in the sample.

- The invention also provides for the storage and retrieval of a collection of target data in a computer data storage apparatus, which can include magnetic disks, optical disks, magneto-optical disks, DRAM, SRAM, SGRAM, SDRAM, RDRAM, DDR RAM, magnetic bubble memory devices, and other data storage devices, including CPU registers and on-CPU data storage arrays. Typically, the target data records are stored as a bit pattern in an array of magnetic domains on a magnetizable medium or as an array of charge states or transistor gate states, such as an array of cells in a DRAM device (e.g., each cell comprised of a transistor and a charge storage area, which may be on the transistor). In one embodiment, the invention provides such storage devices, and computer systems built therewith, comprising a bit pattern encoding a protein expression fingerprint record comprising unique identifiers for at least 10 target data records cross-tabulated with target source.

- When the target is a peptide or nucleic acid, the invention preferably provides a method for identifying related peptide or nucleic acid sequences, comprising performing a computerized comparison between a peptide or nucleic acid sequence assay record stored in or retrieved from a computer storage device or database and at least one other sequence. The

comparison can include a sequence analysis or comparison algorithm or computer program embodiment thereof (e.g., FASTA, TFASTA, GAP, BESTFIT) and/or the comparison may be of the relative amount of a peptide or nucleic acid sequence in a pool of sequences determined from a polypeptide or nucleic acid sample of a specimen.

5 The invention also preferably provides a magnetic disk, such as an IBM-compatible (DOS, Windows, Windows95/98/2000, Windows NT, OS/2) or other format (e.g., Linux, SunOS, Solaris, AIX, SCO Unix, VMS, MV, Macintosh, etc.) floppy diskette or hard (fixed, Winchester) disk drive, comprising a bit pattern encoding data from an assay of the invention in a file format suitable for retrieval and processing in a computerized sequence analysis,  
10 comparison, or relative quantitation method.

The invention also provides a network, comprising a plurality of computing devices linked via a data link, such as an Ethernet cable (coax or 10BaseT), telephone line, ISDN line, wireless network, optical fiber, or other suitable signal transmission medium, whereby at least one network device (e.g., computer, disk array, etc.) comprises a pattern of magnetic  
15 domains (e.g., magnetic disk) and/or charge domains (e.g., an array of DRAM cells) composing a bit pattern encoding data acquired from an assay of the invention.

The invention also provides a method for transmitting assay data that includes generating an electronic signal on an electronic communications device, such as a modem, ISDN terminal adapter, DSL, cable modem, ATM switch, or the like, wherein the signal  
20 includes (in native or encrypted format) a bit pattern encoding data from an assay or a database comprising a plurality of assay results obtained by the method of the invention.

In a preferred embodiment, the invention provides a computer system for comparing a query target to a database containing an array of data structures, such as an assay result obtained by the method of the invention, and ranking database targets based on the degree of  
25 identity and gap weight to the target data. A central processor is preferably initialized to load and execute the computer program for alignment and/or comparison of the assay results. Data for a query target is entered into the central processor via an I/O device. Execution of the computer program results in the central processor retrieving the assay data from the data file, which comprises a binary description of an assay result.

30 The target data or record and the computer program can be transferred to secondary memory, which is typically random access memory (e.g., DRAM, SRAM, SGRAM, or SDRAM). Targets are ranked according to the degree of correspondence between a selected

assay characteristic (e.g., binding to a selected affinity moiety) and the same characteristic of the query target and results are output via an I/O device. For example, a central processor can be a conventional computer (e.g., Intel Pentium, PowerPC, Alpha, PA-8000, SPARC, MIPS 4400, MIPS 10000, VAX, etc.); a program can be a commercial or public domain molecular biology software package (e.g., UWGCG Sequence Analysis Software, Darwin); a data file can be an optical or magnetic disk, a data server, a memory device (e.g., DRAM, SRAM, SGRAM, SDRAM, EPROM, bubble memory, flash memory, etc.); an I/O device can be a terminal comprising a video display and a keyboard, a modem, an ISDN terminal adapter, an Ethernet port, a punched card reader, a magnetic strip reader, or other suitable I/O device.

The invention also preferably provides the use of a computer system, such as that described above, which comprises: (1) a computer; (2) a stored bit pattern encoding a collection of peptide sequence specificity records obtained by the methods of the invention, which may be stored in the computer; (3) a comparison target, such as a query target; and (4) a program for alignment and comparison, typically with rank-ordering of comparison results on the basis of computed similarity values.

# Characteristics of prostate cancer-associated proteins

Prostate cancer proteins of the present invention may be classified as secreted proteins, transmembrane proteins, or intracellular proteins. In one embodiment, the prostate cancer protein is an intracellular protein. Intracellular proteins may be found in the cytoplasm and/or in the nucleus. Intracellular proteins are involved in all aspects of cellular function and replication (including, e.g., signaling pathways); aberrant expression of such proteins often results in unregulated or dysregulated cellular processes (see, e.g., Alberts (ed. 1994) Molecular Biology of the Cell (3d ed.) Garland. For example, many intracellular proteins have enzymatic activity such as protein kinase activity, protein phosphatase activity, protease activity, nucleotide cyclase activity, polymerase activity and the like. Intracellular proteins also serve as docking proteins that are involved in organizing complexes of proteins, or targeting proteins to various subcellular localizations, and are involved in maintaining the structural integrity of organelles.

An increasingly appreciated concept in characterizing proteins is the presence in the proteins of one or more structural motifs for which defined functions have been attributed. In

addition to the highly conserved sequences found in the enzymatic domain of proteins, highly conserved sequences have been identified in proteins that are involved in protein-protein interaction. For example, Src-homology-2 (SH2) domains bind tyrosine-phosphorylated targets in a sequence dependent manner. PTB domains, which are distinct from SH2 domains, also bind tyrosine phosphorylated targets. SH3 domains bind to proline-rich targets. In addition, PH domains, tetratricopeptide repeats and WD domains to name only a few, have been shown to mediate protein-protein interactions. Some of these may also be involved in binding to phospholipids or other second messengers. As will be appreciated by one of ordinary skill in the art, these motifs can be identified on the basis of amino acid sequence; thus, an analysis of the sequence of proteins may provide insight into both the enzymatic potential of the molecule and/or molecules with which the protein may associate. One useful database is Pfam (protein families), which is a large collection of multiple sequence alignments and hidden Markov models covering many common protein domains. Versions are available via the internet from Washington University in St. Louis, the Sanger Center in England, and the Karolinska Institute in Sweden (see, e.g., Bateman, et al. (2000) Nuc. Acids Res. 28:263-266; Sonnhammer, et al. (1997) Proteins 28:405-420; Bateman, et al. (1999) Nuc. Acids Res. 27:260-262; and Sonnhammer, et al. (1998) Nuc. Acids Res. 26:320-322.

In another embodiment, the prostate cancer sequences are transmembrane proteins. Transmembrane proteins are molecules that span a phospholipid bilayer of a cell. They may have an intracellular domain, an extracellular domain, or both. The intracellular domains of such proteins may have a number of functions including those already described for intracellular proteins. For example, the intracellular domain may have enzymatic activity and/or may serve as a binding site for additional proteins. Frequently the intracellular domain of transmembrane proteins serves both roles. For example certain receptor tyrosine kinases have both protein kinase activity and SH2 domains. In addition, autophosphorylation of tyrosines on the receptor molecule itself, creates binding sites for additional SH2 domain containing proteins.

Transmembrane proteins may contain from one to many transmembrane domains. For example, receptor tyrosine kinases, certain cytokine receptors, receptor guanylyl cyclases and receptor serine/threonine protein kinases contain a single transmembrane domain. However, various other proteins including channels and adenylyl cyclases contain numerous



transmembrane domains. Many important cell surface receptors such as G protein coupled receptors (GPCRs) are classified as "seven transmembrane domain" proteins, as they contain 7 membrane spanning regions. Characteristics of transmembrane domains include approximately 17 consecutive hydrophobic amino acids that may be followed by charged amino acids. Therefore, upon analysis of the amino acid sequence of a particular protein, the localization and number of transmembrane domains within the protein may be predicted (see, e.g., PSORT web site <http://psort.nibb.ac.jp/>). Important transmembrane protein receptors include, but are not limited to the insulin receptor, insulin-like growth factor receptor, human growth hormone receptor, glucose transporters, transferrin receptor, epidermal growth factor receptor, low density lipoprotein receptor, epidermal growth factor receptor, leptin receptor, and interleukin receptors, e.g., IL-1 receptor, IL-2 receptor, etc.

The extracellular domains of transmembrane proteins are diverse; however, conserved motifs are found repeatedly among various extracellular domains. Conserved structure and/or functions have been ascribed to different extracellular motifs. Many extracellular domains are involved in binding to other molecules. In one aspect, extracellular domains are found on receptors. Factors that bind the receptor domain include circulating ligands, which may be peptides, proteins, or small molecules such as adenosine and the like. For example, growth factors such as EGF, FGF, and PDGF are circulating growth factors that bind to their cognate receptors to initiate a variety of cellular responses. Other factors include cytokines, mitogenic factors, neurotrophic factors and the like. Extracellular domains also bind to cell-associated molecules. In this respect, they mediate cell-cell interactions. Cell-associated ligands can be tethered to the cell, e.g., via a glycosylphosphatidylinositol (GPI) anchor, or may themselves be transmembrane proteins. Extracellular domains also associate with the extracellular matrix and contribute to the maintenance of the cell structure.

Prostate cancer proteins that are transmembrane are particularly preferred in the present invention as they are readily accessible targets for immunotherapeutics, as are described herein. In addition, as outlined below, transmembrane proteins can be also useful in imaging modalities. Antibodies may be used to label such readily accessible proteins in situ. Alternatively, antibodies can also label intracellular proteins, in which case samples are typically permeabilized to provide access to intracellular proteins. In addition, some membrane proteins can be processed to release a soluble protein, or to expose a residual

fragment. Released soluble proteins may be useful diagnostic markers, processed residual protein fragments may be useful prostate markers of disease.

It will also be appreciated by those in the art that a transmembrane protein can be made soluble by removing transmembrane sequences, e.g., through recombinant methods.

- 5 Furthermore, transmembrane proteins that have been made soluble can be made to be secreted through recombinant means by adding an appropriate signal sequence.

In another embodiment, the prostate cancer proteins are secreted proteins; the secretion of which can be either constitutive or regulated. These proteins may have a signal peptide or signal sequence that targets the molecule to the secretory pathway. Secreted  
10 proteins are involved in numerous physiological events; by virtue of their circulating nature, they often serve to transmit signals to various other cell types. The secreted protein may function in an autocrine manner (acting on the cell that secreted the factor), a paracrine manner (acting on cells in close proximity to the cell that secreted the factor), an endocrine manner (acting on cells at a distance, e.g. secretion into the blood stream), or an exocrine  
15 manner (secretion, e.g., through a duct or to adjacent epithelial surface as sweat glands, sebaceous glands, pancreatic ducts, lacrimal glands, mammary glands, sex producing glands of the ear, etc.). Thus secreted molecules find use in modulating or altering numerous aspects of physiology. Prostate cancer proteins that are secreted proteins are particularly preferred in the present invention as they serve as good targets for diagnostic markers, e.g., for blood,  
20 plasma, serum, or stool tests. Those which are enzymes may be antibody or small molecule targets. Others may be useful as vaccine targets, e.g., via CTL mechanisms.

#### Use of prostate cancer nucleic acids

- As described above, prostate cancer sequence is initially identified by substantial  
25 nucleic acid and/or amino acid sequence homology or linkage to the prostate cancer sequences outlined herein. Such homology can be based upon the overall nucleic acid or amino acid sequence, and is generally determined as outlined below, using either homology programs or hybridization conditions. Typically, linked sequences on a mRNA are found on the same molecule.

- 30 The prostate cancer nucleic acid sequences of the invention, e.g., the sequences in Tables 1A-4, can be fragments of larger genes, i.e., they are nucleic acid segments. "Genes" in this context includes coding regions, non-coding regions, and mixtures of coding and non-

coding regions. Accordingly, as will be appreciated by those in the art, using the sequences provided herein, extended sequences, in either direction, of the prostate cancer genes can be obtained, using techniques well known in the art for cloning either longer sequences or the full length sequences; see Ausubel, et al., supra. Much can be done by informatics and many sequences can be clustered to include multiple sequences corresponding to a single gene, e.g., systems such as UniGene (see, <http://www.ncbi.nlm.nih.gov/UniGene/>).

Once the prostate cancer nucleic acid is identified, it can be cloned and, if necessary, its constituent parts recombined to form the entire prostate cancer nucleic acid coding regions or the entire mRNA sequence. Once isolated from its natural source, e.g., contained within a plasmid or other vector or excised therefrom as a linear nucleic acid segment, the recombinant prostate cancer nucleic acid can be further-used as a probe to identify and isolate other prostate cancer nucleic acids, e.g., extended coding regions. It can also be used as a "precursor" nucleic acid to make modified or variant prostate cancer nucleic acids and proteins.

The prostate cancer nucleic acids of the present invention are used in several ways. In a first embodiment, nucleic acid probes to the prostate cancer nucleic acids are made and attached to biochips to be used in screening and diagnostic methods, as outlined below, or for administration, e.g., for gene therapy, vaccine, and/or antisense applications. Alternatively, the prostate cancer nucleic acids that include coding regions of prostate cancer proteins can be put into expression vectors for the expression of prostate cancer proteins, again for screening purposes or for administration to a patient.

In a preferred embodiment, nucleic acid probes to prostate cancer nucleic acids (both the nucleic acid sequences outlined in the figures and/or the complements thereof) are made. The nucleic acid probes attached to the biochip are designed to be substantially complementary to the prostate cancer nucleic acids, i.e., the target sequence (either the target sequence of the sample or to other probe sequences, e.g., in sandwich assays), such that hybridization of the target sequence and the probes of the present invention occurs. As outlined below, this complementarity need not be perfect; there may be base pair mismatches which will interfere with hybridization between the target sequence and the single stranded nucleic acids of the present invention. However, if the number of mutations is so great that no hybridization can occur under even the least stringent of hybridization conditions, the sequence is not a complementary target sequence. Thus, by "substantially complementary"

herein is meant that the probes are sufficiently complementary to the target sequences to hybridize under normal reaction conditions, particularly high stringency conditions, as outlined herein.

A nucleic acid probe is generally single stranded but can be partially single and partially double stranded. The strandedness of the probe is dictated by the structure, composition, and properties of the target sequence. In general, the nucleic acid probes range from about 8 to about 100 bases long, with from about 10 to about 80 bases being preferred, and from about 30 to about 50 bases being particularly preferred. That is, generally whole genes are not used. In some embodiments, much longer nucleic acids can be used, up to hundreds of bases.

In a preferred embodiment, more than one probe per sequence is used, with either overlapping probes or probes to different sections of the target being used. That is, two, three, four or more probes, with three being preferred, are used to build in a redundancy for a particular target. The probes can be overlapping (i.e., have some sequence in common), or separate. In some cases, PCR primers may be used to amplify signal for higher sensitivity.

As will be appreciated by those in the art, nucleic acids can be attached or immobilized to a solid support in a wide variety of ways. By "immobilized" and grammatical equivalents herein is meant the association or binding between the nucleic acid probe and the solid support is sufficient to be stable under the conditions of binding, washing, analysis, and removal as outlined below. The binding can typically be covalent or non-covalent. By "non-covalent binding" and grammatical equivalents herein is meant one or more of electrostatic, hydrophilic, and hydrophobic interactions. Included in non-covalent binding is the covalent attachment of a molecule, such as, streptavidin to the support and the non-covalent binding of the biotinylated probe to the streptavidin. By "covalent binding" and grammatical equivalents herein is meant that the two moieties, the solid support and the probe, are attached by at least one bond, including sigma bonds, pi bonds and coordination bonds. Covalent bonds can be formed directly between the probe and the solid support or can be formed by a cross linker or by inclusion of a specific reactive group on either the solid support or the probe or both molecules. Immobilization may also involve a combination of covalent and non-covalent interactions.

In general, the probes are attached to the biochip in a wide variety of ways, as will be appreciated by those in the art. As described herein, the nucleic acids can either be

synthesized first, with subsequent attachment to the biochip, or can be directly synthesized on the biochip.

The biochip comprises a suitable solid substrate. By "substrate" or "solid support" or other grammatical equivalents herein is meant a material that can be modified to contain discrete individual sites appropriate for the attachment or association of the nucleic acid probes and is amenable to at least one detection method. As will be appreciated by those in the art, the number of possible substrates are very large, and include, but are not limited to, glass and modified or functionalized glass, plastics (including acrylics, polystyrene and copolymers of styrene and other materials, polypropylene, polyethylene, polybutylene, polyurethanes, Teflon, etc.), polysaccharides, nylon or nitrocellulose, resins, silica or silica-based materials including silicon and modified silicon, carbon, metals, inorganic glasses, plastics, etc. In general, the substrates allow optical detection and do not appreciably fluoresce. A preferred substrate is described in WO0055627, herein incorporated by reference in its entirety.

Generally the substrate is planar, although as will be appreciated by those in the art, other configurations of substrates may be used as well. For example, the probes may be placed on the inside surface of a tube, for flow-through sample analysis to minimize sample volume. Similarly, the substrate may be flexible, such as a flexible foam, including closed cell foams made of particular plastics.

In a preferred embodiment, the surface of the biochip and the probe may be derivatized with chemical functional groups for subsequent attachment of the two. Thus, e.g., the biochip is derivatized with a chemical functional group including, but not limited to, amino groups, carboxy groups, oxo groups and thiol groups, with amino groups being particularly preferred. Using these functional groups, the probes can be attached using functional groups on the probes. For example, nucleic acids containing amino groups can be attached to surfaces comprising amino groups, e.g., using linkers as are known in the art; e.g., homo- or hetero-bifunctional linkers as are well known (see 1994 Pierce Chemical Company catalog, technical section on cross-linkers, pages 155-200). In addition, in some cases, additional linkers, such as alkyl groups (including substituted and heteroalkyl groups) may be used.

In this embodiment, oligonucleotides are synthesized as is known in the art, and then attached to the surface of the solid support. As will be appreciated by those skilled in the art,

either the 5' or 3' terminus may be attached to the solid support, or attachment may be via an internal nucleoside.

In another embodiment, the immobilization to the solid support may be very strong, yet non-covalent. For example, biotinylated oligonucleotides can be made, which bind to surfaces covalently coated with streptavidin, resulting in attachment.

Alternatively, the oligonucleotides may be synthesized on the surface, as is known in the art. For example, photoactivation techniques utilizing photopolymerization compounds and techniques are used. In a preferred embodiment, the nucleic acids can be synthesized in situ, using well known photolithographic techniques, such as those described in WO 95/25116; WO 95/35505; U.S. Patent Nos. 5,700,637 and 5,445,934; and references cited within, all of which are expressly incorporated by reference; these methods of attachment form the basis of the Affymetrix GeneChip™ technology.

Often, amplification-based assays are performed to measure the expression level of prostate cancer-associated sequences. These assays are typically performed in conjunction with reverse transcription. In such assays, a prostate cancer-associated nucleic acid sequence acts as a template in an amplification reaction (e.g., Polymerase Chain Reaction, or PCR). In a quantitative amplification, the amount of amplification product will be proportional to the amount of template in the original sample. Comparison to appropriate controls provides a measure of the amount of prostate cancer-associated RNA. Methods of quantitative amplification are well known to those of skill in the art. Detailed protocols for quantitative PCR are provided, e.g., in Innis, et al. (1990) PCR Protocols: A Guide to Methods and Applications Academic Press.

In some embodiments, a TaqMan based assay is used to measure expression. TaqMan based assays use a fluorogenic oligonucleotide probe that contains a 5' fluorescent dye and a 3' quenching agent. The probe hybridizes to a PCR product, but cannot itself be extended due to a blocking agent at the 3' end. When the PCR product is amplified in subsequent cycles, the 5' nuclease activity of the polymerase, e.g., AmpliTaq, results in the cleavage of the TaqMan probe. This cleavage separates the 5' fluorescent dye and the 3' quenching agent, thereby resulting in an increase in fluorescence as a function of amplification (see, e.g., literature provided by Perkin-Elmer, e.g., www2.perkin-elmer.com).

Other suitable amplification methods include, but are not limited to, ligase chain reaction (LCR) (see Wu and Wallace (1989) Genomics 4:560-569, Landegren, et al. (1988)

Science 241:1077-1080, and Barringer, et al. (1990) *Gene* 89:117-122), transcription amplification (Kwoh, et al. (1989) *Proc. Natl. Acad. Sci. USA* 86:1173-1177), self-sustained sequence replication (Guatelli, et al. (1990) *Proc. Nat. Acad. Sci. USA* 87:1874-1878), dot PCR, and linker adapter PCR, etc.

5

#### Expression of prostate cancer proteins from nucleic acids

In a preferred embodiment, prostate cancer nucleic acids, e.g., encoding prostate cancer proteins are used to make a variety of expression vectors to express prostate cancer proteins which can then be used in screening assays, as described below. Expression vectors and recombinant DNA technology are well known to those of skill in the art (see, e.g., Ausubel, supra, and Fernandez and Hoeffler (eds. 1999) *Gene Expression Systems* Academic Press) and are used to express proteins. The expression vectors may be either self-replicating extrachromosomal vectors or vectors which integrate into a host genome. Generally, these expression vectors include transcriptional and translational regulatory nucleic acid operably linked to the nucleic acid encoding the prostate cancer protein. The term "control sequences" refers to DNA sequences used for the expression of an operably linked coding sequence in a particular host organism. Control sequences that are suitable for prokaryotes, e.g., include a promoter, optionally an operator sequence, and a ribosome binding site. Eukaryotic cells are known to utilize promoters, polyadenylation signals, and enhancers.

Nucleic acid is "operably linked" when it is placed into a functional relationship with another nucleic acid sequence. For example, DNA for a presequence or secretory leader is operably linked to DNA for a polypeptide if it is expressed as a preprotein that participates in the secretion of the polypeptide; a promoter or enhancer is operably linked to a coding sequence if it affects the transcription of the sequence; a ribosome binding site is operably linked to a coding sequence if it is positioned so as to facilitate translation, and sequences may be operably linked when they are physically linked on the same molecule. Generally, "operably linked" means that the DNA sequences being linked are contiguous, and, in the case of a secretory leader, contiguous and in reading phase. However, enhancers do not have to be contiguous. Linking is typically accomplished by ligation at convenient restriction sites. If such sites do not exist, synthetic oligonucleotide adaptors or linkers are used in accordance with conventional practice. Transcriptional and translational regulatory nucleic acid will generally be appropriate to the host cell used to express the prostate cancer protein.

Numerous types of appropriate expression vectors, and suitable regulatory sequences are known in the art for a variety of host cells.

In general, transcriptional and translational regulatory sequences may include, but are not limited to, promoter sequences, ribosomal binding sites, transcriptional start and stop sequences, translational start and stop sequences, and enhancer or activator sequences. In a preferred embodiment, the regulatory sequences include a promoter and transcriptional start and stop sequences.

Promoter sequences encode either constitutive or inducible promoters. The promoters may be either naturally occurring promoters or hybrid promoters. Hybrid promoters, which combine elements of more than one promoter, are also known in the art, and are useful in the present invention.

In addition, an expression vector may comprise additional elements. For example, the expression vector may have two replication systems, thus allowing it to be maintained in two organisms, e.g., in mammalian or insect cells for expression and in a prokaryotic host for cloning and amplification. Furthermore, for integrating expression vectors, the expression vector contains at least one sequence homologous to the host cell genome, and preferably two homologous sequences which flank the expression construct. The integrating vector may be directed to a specific locus in the host cell by selecting the appropriate homologous sequence for inclusion in the vector. Constructs for integrating vectors are well known in the art (e.g., Fernandez and Hoeffler, supra).

In addition, in a preferred embodiment, the expression vector contains a selectable marker gene to allow the selection of transformed host cells. Selection genes are well known in the art and will vary with the host cell used.

The prostate cancer proteins of the present invention are produced by culturing a host cell transformed with an expression vector containing nucleic acid encoding a prostate cancer protein, under the appropriate conditions to induce or cause expression of the prostate cancer protein. Conditions appropriate for prostate cancer protein expression will vary with the choice of the expression vector and the host cell, and will be easily ascertained by one skilled in the art through routine experimentation or optimization. For example, the use of constitutive promoters in the expression vector will require optimizing the growth and proliferation of the host cell, while the use of an inducible promoter requires the appropriate growth conditions for induction. In addition, in some embodiments, the timing of the harvest



is important. For example, the baculoviral systems used in insect cell expression are lytic viruses, and thus harvest time selection can be crucial for product yield.

Appropriate host cells include yeast, bacteria, archaeobacteria, fungi, and insect and animal cells, including mammalian cells. Of particular interest are *Saccharomyces cerevisiae* and other yeasts, *E. coli*, *Bacillus subtilis*, Sf9 cells, C129 cells, 293 cells, *Neurospora*, BHK, CHO, COS, HeLa cells, HUVEC (human umbilical vein endothelial cells), THP1 cells (a macrophage cell line) and various other human cells and cell lines.

In a preferred embodiment, the prostate cancer proteins are expressed in mammalian cells. Mammalian expression systems are also known in the art, and include retroviral and adenoviral systems. One expression vector system is a retroviral vector system such as is generally described in PCT/US97/01019 and PCT/US97/01048, both of which are hereby expressly incorporated by reference. Of particular use as mammalian promoters are the promoters from mammalian viral genes, since the viral genes are often highly expressed and have a broad host range. Examples include the SV40 early promoter, mouse mammary tumor virus LTR promoter, adenovirus major late promoter, herpes simplex virus promoter, and the CMV promoter (see, e.g., Fernandez and Hoefler, supra). Typically, transcription termination and polyadenylation sequences recognized by mammalian cells are regulatory regions located 3' to the translation stop codon and thus, together with the promoter elements, flank the coding sequence. Examples of transcription terminator and polyadenylation signals include those derived from SV40.

The methods of introducing exogenous nucleic acid into mammalian hosts, as well as other hosts, is well known in the art, and will vary with the host cell used. Techniques include dextran-mediated transfection, calcium phosphate precipitation, polybrene mediated transfection, protoplast fusion, electroporation, viral infection, encapsulation of the polynucleotide(s) in liposomes, and direct microinjection of the DNA into nuclei.

In a preferred embodiment, prostate cancer proteins are expressed in bacterial systems. Bacterial expression systems are well known in the art. Promoters from bacteriophage may also be used and are known in the art. In addition, synthetic promoters and hybrid promoters are also useful; e.g., the tac promoter is a hybrid of the trp and lac promoter sequences. Furthermore, a bacterial promoter can include naturally occurring promoters of non-bacterial origin that have the ability to bind bacterial RNA polymerase and initiate transcription. In addition to a functioning promoter sequence, an efficient ribosome

binding site is desirable. The expression vector may also include a signal peptide sequence that provides for secretion of the prostate cancer protein in bacteria. The protein is either secreted into the growth media (gram-positive bacteria) or into the periplasmic space, located between the inner and outer membrane of the cell (gram-negative bacteria). The bacterial expression vector may also include a selectable marker gene to allow for the selection of bacterial strains that have been transformed. Suitable selection genes include genes which render the bacteria resistant to drugs such as ampicillin, chloramphenicol, erythromycin, kanamycin, neomycin and tetracycline. Selectable markers also include biosynthetic genes, such as those in the histidine, tryptophan and leucine biosynthetic pathways. These components are assembled into expression vectors. Expression vectors for bacteria are well known in the art, and include vectors for *Bacillus subtilis*, *E. coli*, *Streptococcus cremoris*, and *Streptococcus lividans*, among others (e.g., Fernandez and Hoeffler, supra). The bacterial expression vectors are transformed into bacterial host cells using techniques well known in the art, such as calcium chloride treatment, electroporation, and others.

In one embodiment, prostate cancer proteins are produced in insect cells. Expression vectors for the transformation of insect cells, and in particular, baculovirus-based expression vectors, are well known in the art.

In a preferred embodiment, prostate cancer protein is produced in yeast cells. Yeast expression systems are well known in the art, and include expression vectors for *Saccharomyces cerevisiae*, *Candida albicans* and *C. maltosa*, *Hansenula polymorpha*, *Kluyveromyces fragilis* and *K. lactis*, *Pichia guilliermondii* and *P. pastoris*, *Schizosaccharomyces pombe*, and *Yarrowia lipolytica*.

The prostate cancer protein may also be made as a fusion protein, using techniques well known in the art. Thus, e.g., for the creation of monoclonal antibodies, if the desired epitope is small, the prostate cancer protein may be fused to a carrier protein to form an immunogen. Alternatively, the prostate cancer protein may be made as a fusion protein to increase expression, or for other reasons. For example, when the prostate cancer protein is a prostate cancer peptide, the nucleic acid encoding the peptide may be linked to other nucleic acid for expression purposes.

In a preferred embodiment, the prostate cancer protein is purified or isolated after expression. Prostate cancer proteins may be isolated or purified in a variety of ways known to those skilled in the art depending on what other components are present in the sample.

Standard purification methods include electrophoretic, molecular, immunological and chromatographic techniques, including ion exchange, hydrophobic, affinity, and reverse-phase HPLC chromatography, and chromatofocusing. For example, the prostate cancer protein may be purified using a standard anti-prostate cancer protein antibody column.

5 Ultrafiltration and diafiltration techniques, in conjunction with protein concentration, are also useful. For general guidance in suitable purification techniques, see Scopes (1982) Protein Purification Springer-Verlag. The degree of purification necessary will vary depending on the use of the prostate cancer protein. In some instances no purification will be necessary.

10 Once expressed and purified if necessary, the prostate cancer proteins and nucleic acids are useful in a number of applications. They may be used as immunoselection reagents, as vaccine reagents, as screening agents, etc.

#### Variants of prostate cancer proteins

15 In one embodiment, the prostate cancer proteins are derivative or variant prostate cancer proteins as compared to the wild-type sequence. That is, as outlined more fully below, the derivative prostate cancer peptide will often contain at least one amino acid substitution, deletion or insertion, with amino acid substitutions being particularly preferred. The amino acid substitution, insertion, or deletion may occur at most any residue within the prostate cancer peptide.

20 Also included within one embodiment of prostate cancer proteins of the present invention are amino acid sequence variants. These variants typically fall into one or more of three classes: substitutional, insertional, or deletional variants. These variants ordinarily are prepared by site specific mutagenesis of nucleotides in the DNA encoding the prostate cancer protein, using cassette or PCR mutagenesis or other techniques well known in the art, to  
25 produce DNA encoding the variant, and thereafter expressing the DNA in recombinant cell culture as outlined above. However, variant prostate cancer protein fragments having up to about 100-150 residues may be prepared by in vitro synthesis using established techniques. Amino acid sequence variants are characterized by the predetermined nature of the variation, a feature that sets them apart from naturally occurring allelic or interspecies variation of the  
30 prostate cancer protein amino acid sequence. The variants typically exhibit the same qualitative biological activity as the naturally occurring analogue, although variants can also be selected which have modified characteristics as will be more fully outlined below.

While the site or region for introducing an amino acid sequence variation is predetermined, the mutation per se need not be predetermined. For example, in order to optimize the performance of a mutation at a given site, random mutagenesis may be conducted at the target codon or region and the expressed prostate cancer variants screened for the optimal combination of desired activity. Techniques for making substitution mutations at predetermined sites in DNA having a known sequence are well known, e.g., M13 primer mutagenesis and PCR mutagenesis. Screening of the mutants is done using assays of prostate cancer protein activities.

Amino acid substitutions are typically of single residues; insertions usually will be on the order of from about 1 to 20 amino acids, although considerably larger insertions may be tolerated. Deletions range from about 1 to about 20 residues, although in some cases deletions may be much larger.

Substitutions, deletions, insertions or a combination thereof may be used to arrive at a final derivative. Generally these changes are done on a few amino acids to minimize the alteration of the molecule. However, larger changes may be tolerated in certain circumstances. When small alterations in the characteristics of the prostate cancer protein are desired, substitutions are generally made in accordance with the amino acid substitution relationships provided in the definition section.

The variants typically exhibit the same qualitative biological activity and will elicit the same immune response as the naturally-occurring analog, although variants also are selected to modify the characteristics of the prostate cancer proteins as needed. Alternatively, the variant may be designed such that the biological activity of the prostate cancer protein is altered. For example, glycosylation sites may be altered or removed.

Substantial changes in function or immunological identity are made by selecting substitutions that are less conservative than those described above. For example, substitutions may be made which more significantly affect: the structure of the polypeptide backbone in the area of the alteration, for example the alpha-helical or beta-sheet structure; the charge or hydrophobicity of the molecule at the target site; or the bulk of the side chain. The substitutions which in general are expected to produce the greatest changes in the polypeptide's properties are those in which (a) a hydrophilic residue, e.g., serinyl or thronyl is substituted for (or by) a hydrophobic residue, e.g., leucyl, isoleucyl, phenylalanyl, valyl or alanyl; (b) a cysteine or proline is substituted for (or by) another residue; (c) a residue having

an electropositive side chain, e.g., lysyl, arginyl, or histidyl, is substituted for (or by) an electronegative residue, e.g., glutamyl or aspartyl; or (d) a residue having a bulky side chain, e.g., phenylalanine, is substituted for (or by) one not having a side chain, e.g., glycine.

Covalent modifications of prostate cancer polypeptides are included within the scope of this invention. One type of covalent modification includes reacting targeted amino acid residues of a prostate cancer polypeptide with an organic derivatizing agent that is capable of reacting with selected side chains or the N-or C-terminal residues of a prostate cancer polypeptide. Derivatization with bifunctional agents is useful, for instance, for crosslinking prostate cancer polypeptides to a water-insoluble support matrix or surface for use in the method for purifying anti-prostate cancer polypeptide antibodies or screening assays, as is more fully described below. Commonly used crosslinking agents include, e.g., 1,1-bis(diazoacetyl)-2-phenylethane, glutaraldehyde, N-hydroxysuccinimide esters, e.g., esters with 4-azidosalicylic acid, homobifunctional imidoesters, including disuccinimidyl esters such as 3,3'-dithiobis(succinimidylpropionate), bifunctional maleimides such as bis-N-maleimido-1,8-octane and agents such as methyl-3-((p-azidophenyl)dithio)propioimide.

Other modifications include deamidation of glutaminyl and asparaginyl residues to the corresponding glutamyl and aspartyl residues, respectively, hydroxylation of proline and lysine, phosphorylation of hydroxyl groups of serinyl, threonyl or tyrosyl residues, methylation of the amino groups of the lysine, arginine, and histidine side chains (e.g., pp. 79-86, Creighton (1983) Proteins: Structure and Molecular Properties Freeman), acetylation of the N-terminal amine, and amidation of a C-terminal carboxyl group.

Another type of covalent modification of the prostate cancer polypeptide included within the scope of this invention comprises altering the native glycosylation pattern of the polypeptide. "Altering the native glycosylation pattern" is intended for purposes herein to mean deleting one or more carbohydrate moieties found in native sequence prostate cancer polypeptide, and/or adding one or more glycosylation sites that are not present in the native sequence prostate cancer polypeptide. Glycosylation patterns can be altered in many ways. For example the use of different cell types to express prostate cancer-associated sequences can result in different glycosylation patterns.

Addition of glycosylation sites to prostate cancer polypeptides may also be accomplished by altering the amino acid sequence thereof. The alteration may be made, e.g., by the addition of, or substitution by, one or more serine or threonine residues to the native

sequence prostate cancer polypeptide (for O-linked glycosylation sites). The prostate cancer amino acid sequence may optionally be altered through changes at the DNA level, particularly by mutating the DNA encoding the prostate cancer polypeptide at preselected bases such that codons are generated that will translate into the desired amino acids.

Another means of increasing the number of carbohydrate moieties on the prostate cancer polypeptide is by chemical or enzymatic coupling of glycosides to the polypeptide. Such methods are described in the art, e.g., in WO 87/05330, and pp. 259-306 in Aplin and Wriston (1981) CRC Crit. Rev. Biochem.

Removal of carbohydrate moieties present on the prostate cancer polypeptide may be accomplished chemically or enzymatically or by mutational substitution of codons encoding for amino acid residues that serve as targets for glycosylation. Chemical deglycosylation techniques are known in the art and described, e.g., by Hakimuddin, et al. (1987) Arch. Biochem. Biophys. 259:52-57; and Edge, et al. (1981) Anal. Biochem. 118:131-137. Enzymatic cleavage of carbohydrate moieties on polypeptides can be achieved by the use of a variety of endo- and exo-glycosidases as described by Thotakura, et al. (1987) Meth. Enzymol. 138:350-359.

Another type of covalent modification of prostate cancer comprises linking the prostate cancer polypeptide to one of a variety of nonproteinaceous polymers, e.g., polyethylene glycol, polypropylene glycol, or polyoxyalkylenes, in the manner set forth in U.S. Patent Nos. 4,640,835; 4,496,689; 4,301,144; 4,670,417; 4,791,192; or 4,179,337.

Prostate cancer polypeptides of the present invention may also be modified in a way to form chimeric molecules comprising a prostate cancer polypeptide fused to another, heterologous polypeptide or amino acid sequence. In one embodiment, such a chimeric molecule comprises a fusion of a prostate cancer polypeptide with a tag polypeptide which provides an epitope to which an anti-tag antibody can selectively bind. The epitope tag is generally placed at the amino- or carboxyl-terminus of the prostate cancer polypeptide. The presence of such epitope-tagged forms of a prostate cancer polypeptide can be detected using an antibody against the tag polypeptide. Also, provision of the epitope tag enables the prostate cancer polypeptide to be readily purified by affinity purification using an anti-tag antibody or another type of affinity matrix that binds to the epitope tag. In an alternative embodiment, the chimeric molecule may comprise a fusion of a prostate cancer polypeptide

with an immunoglobulin or a particular region of an immunoglobulin. For a bivalent form of the chimeric molecule, such a fusion could be to the Fc region of an IgG molecule.

Various tag polypeptides and their respective antibodies are well known in the art.

Examples include poly-histidine (poly-his) or poly-histidine-glycine (poly-his-gly) tags; HIS6 and metal chelation tags, the flu HA tag polypeptide and its antibody 12CA5 (Field, et al. (1988) Mol. Cell. Biol. 8:2159-2165; the c-myc tag and the 8F9, 3C7, 6E10, G4, B7, and 9E10 antibodies thereto (Evan, et al. (1985) Molecular and Cellular Biology 5:3610-3616); and the Herpes Simplex virus glycoprotein D (gD) tag and its antibody (Paborsky, et al. (1990) Protein Engineering 3:547-553). Other tag polypeptides include the Flag-peptide (Hopp, et al. (1988) BioTechnology 6:1204-1210); the KT3 epitope peptide (Martin, et al. (1992) Science 255:192-194); tubulin epitope peptide (Skinner, et al. (1991) J. Biol. Chem. 266:15163-15166); and the T7 gene 10 protein peptide tag (Lutz-Freyermuth, et al. (1990) Proc. Natl. Acad. Sci. USA 87:6393-6397).

Also included are other prostate cancer proteins of the prostate cancer family, and prostate cancer proteins from other organisms, which are cloned and expressed as outlined below. Thus, probe or degenerate polymerase chain reaction (PCR) primer sequences may be used to find other related prostate cancer proteins from humans or other organisms. As will be appreciated by those in the art, particularly useful probe and/or PCR primer sequences include the unique areas of the prostate cancer nucleic acid sequence. As is generally known in the art, preferred PCR primers are from about 15 to about 35 nucleotides in length, with from about 20 to about 30 being preferred, and may contain inosine as needed. The conditions for the PCR reaction are well known in the art (e.g., Innis, PCR Protocols, supra).

#### Antibodies to prostate cancer proteins

In a preferred embodiment, when the prostate cancer protein is to be used to generate antibodies, e.g., for immunotherapy or immunodiagnosis, the prostate cancer protein should share at least one epitope or determinant with the full length protein. By "epitope" or "determinant" herein is typically meant a portion of a protein which will generate and/or bind an antibody or T-cell receptor in the context of MHC. Thus, in most instances, antibodies made to a smaller prostate cancer protein will be able to bind to the full-length protein, particularly linear epitopes. In a preferred embodiment, the epitope is unique; that is, antibodies generated to a unique epitope show little or no cross-reactivity.

Methods of preparing polyclonal antibodies are known to the skilled artisan (e.g., Coligan, supra; and Harlow and Lane, supra). Polyclonal antibodies can be raised in a mammal, e.g., by one or more injections of an immunizing agent and, if desired, an adjuvant. Typically, the immunizing agent and/or adjuvant will be injected in the mammal by multiple subcutaneous or intraperitoneal injections. The immunizing agent may include a protein encoded by a nucleic acid of the figures or fragment thereof or a fusion protein thereof. It may be useful to conjugate the immunizing agent to a protein known to be immunogenic in the mammal being immunized. Examples of such immunogenic proteins include but are not limited to keyhole limpet hemocyanin, serum albumin, bovine thyroglobulin, and soybean trypsin inhibitor. Examples of adjuvants which may be employed include Freund's complete adjuvant and MPL-TDM adjuvant (monophosphoryl Lipid A, synthetic trehalose dicorynomycolate). The immunization protocol may be selected by one skilled in the art without undue experimentation.

The antibodies may, alternatively, be monoclonal antibodies. Monoclonal antibodies may be prepared using hybridoma methods, such as those described by Kohler and Milstein (1975) *Nature* 256:495. In a hybridoma method, a mouse, hamster, or other appropriate host animal, is typically immunized with an immunizing agent to elicit lymphocytes that produce or are capable of producing antibodies that will specifically bind to the immunizing agent. Alternatively, the lymphocytes may be immunized in vitro. The immunizing agent will typically include a polypeptide encoded by a nucleic acid of Tables 1A-4 or fragment thereof, or a fusion protein thereof. Generally, either peripheral blood lymphocytes ("PBLs") are used if cells of human origin are desired, or spleen cells or lymph node cells are used if non-human mammalian sources are desired. The lymphocytes are then fused with an immortalized cell line using a suitable fusing agent, such as polyethylene glycol, to form a hybridoma cell (see pp. 59-103 in Goding (1986) Monoclonal Antibodies: Principles and Practice Academic Press). Immortalized cell lines are usually transformed mammalian cells, particularly myeloma cells of rodent, bovine and human origin. Usually, rat or mouse myeloma cell lines are employed. The hybridoma cells may be cultured in a suitable culture medium that preferably contains one or more substances that inhibit the growth or survival of the unfused, immortalized cells. For example, if the parental cells lack the enzyme hypoxanthine guanine phosphoribosyl transferase (HGPRT or HPRT), the culture medium



for the hybridomas typically will include hypoxanthine, aminopterin, and thymidine ("HAT medium"), which substances prevent the growth of HGPRT-deficient cells.

In one embodiment, the antibodies are bispecific antibodies. Bispecific antibodies are monoclonal, preferably human or humanized, antibodies that have binding specificities for at least two different antigens or that have binding specificities for two epitopes on the same antigen. In one embodiment, one of the binding specificities is for a protein encoded by a nucleic acid of Tables 1A-4 or a fragment thereof, the other one is for another antigen, and preferably for a cell-surface protein or receptor or receptor subunit, preferably one that is tumor specific. Alternatively, tetramer-type technology may create multivalent reagents.

In a preferred embodiment, the antibodies to prostate cancer protein are capable of reducing or eliminating a biological function of a prostate cancer protein, as is described below. That is, the addition of anti-prostate cancer protein antibodies (either polyclonal or preferably monoclonal) to prostate cancer tissue (or cells containing prostate cancer) may reduce or eliminate the prostate cancer. Generally, at least a 25% decrease in activity, growth, size or the like is preferred, with at least about 50% being particularly preferred and about a 95-100% decrease being especially preferred.

In a preferred embodiment the antibodies to the prostate cancer proteins are humanized antibodies (e.g., Xenerex Biosciences; Medarex, Inc.; Abgenix, Inc.; Protein Design Labs, Inc.). Humanized forms of non-human (e.g., murine) antibodies are chimeric molecules of immunoglobulins, immunoglobulin chains or fragments thereof (such as Fv, Fab, Fab', F(ab')<sub>2</sub> or other antigen-binding subsequences of antibodies) which contain minimal sequence derived from non-human immunoglobulin. Humanized antibodies include human immunoglobulins (recipient antibody) in which residues from a complementary determining region (CDR) of the recipient are replaced by residues from a CDR of a non-human species (donor antibody) such as mouse, rat or rabbit having the desired specificity, affinity and capacity. In some instances, Fv framework residues of the human immunoglobulin are replaced by corresponding non-human residues. Humanized antibodies may also comprise residues which are found neither in the recipient antibody nor in the imported CDR or framework sequences. In general, a humanized antibody will comprise substantially all of at least one, and typically two, variable domains, in which all or substantially all of the CDR regions correspond to those of a non-human immunoglobulin and all or substantially all of the framework (FR) regions are those of a human



art, the antigen may be provided by injecting a polypeptide against which antibodies are desired to be raised into a recipient, or contacting the recipient with a nucleic acid capable of expressing the antigen and under conditions for expression of the antigen, leading to an immune response.

5 In a preferred embodiment the prostate cancer proteins against which antibodies are raised are secreted proteins as described above. Without being bound by theory, antibodies used for treatment, bind and prevent the secreted protein from binding to its receptor, thereby inactivating the secreted prostate cancer protein.

10 In another preferred embodiment, the prostate cancer protein to which antibodies are raised is a transmembrane protein. Without being bound by theory, antibodies used for treatment bind the extracellular domain of the prostate cancer protein and prevent it from binding to other proteins, such as circulating ligands or cell-associated molecules. The antibody may cause down-regulation of the transmembrane prostate cancer protein. As will be appreciated by one of ordinary skill in the art, the antibody may be a competitive, non-competitive or uncompetitive inhibitor of protein binding to the extracellular domain of the prostate cancer protein. The antibody is also often an antagonist of the prostate cancer protein. Further, the antibody may prevent activation of the transmembrane prostate cancer protein. In one aspect, when the antibody prevents the binding of other molecules to the prostate cancer protein, the antibody prevents growth of the cell. The antibody may also be used to target or sensitize the cell to cytotoxic agents, including, but not limited to TNF- $\alpha$ , TNF- $\beta$ , IL-1, INF- $\gamma$ , and IL-2, or chemotherapeutic agents including 5FU, vinblastine, actinomycin D, cisplatin, methotrexate, and the like. In some instances the antibody belongs to a sub-type that activates serum complement when complexed with the transmembrane protein thereby mediating cytotoxicity or antigen-dependent cytotoxicity (ADCC). Thus, 20 prostate cancer is treated by administering to a patient antibodies directed against the transmembrane prostate cancer protein. Antibody-labeling may activate a co-toxin, localize a toxin payload, or otherwise provide means to locally ablate cells.

25 In another preferred embodiment, the antibody is conjugated to an effector moiety. The effector moiety can be a labeling moiety such as a radioactive label or fluorescent label, or can be a therapeutic moiety. In one aspect the therapeutic moiety is a small molecule that modulates the activity of the prostate cancer protein. In another aspect the therapeutic moiety modulates the activity of molecules associated with or in close proximity to the prostate

cancer protein. The therapeutic moiety may inhibit enzymatic activity such as protease or collagenase or protein kinase activity associated with prostate cancer.

In a preferred embodiment, the therapeutic moiety can also be a cytotoxic agent. In this method, targeting the cytotoxic agent to prostate cancer tissue or cells, results in a reduction in the number of afflicted cells, thereby reducing symptoms associated with prostate cancer. Cytotoxic agents are numerous and varied and include, but are not limited to, cytotoxic drugs or toxins or active fragments of such toxins. Suitable toxins and their corresponding fragments include diphtheria A chain, exotoxin A chain, ricin A chain, abrin A chain, curcin, crotin, phenomycin, enomycin, saporin, auristatin, and the like. Cytotoxic agents also include radiochemicals made by conjugating radioisotopes to antibodies raised against prostate cancer proteins, or binding of a radionuclide to a chelating agent that has been covalently attached to the antibody. Targeting the therapeutic moiety to transmembrane prostate cancer proteins not only serves to increase the local concentration of therapeutic moiety in the prostate cancer afflicted area, but also serves to reduce deleterious side effects, e.g., by binding to normal tissues, that may be associated with the therapeutic moiety.

In another preferred embodiment, the prostate cancer protein against which the antibodies are raised is an intracellular protein. In this case, the antibody may be conjugated to a protein which facilitates entry into the cell. In one case, the antibody enters the cell by endocytosis. In another embodiment, a nucleic acid encoding the antibody is administered to the individual or cell. Moreover, wherein the prostate cancer protein can be targeted within a cell, i.e., the nucleus, an antibody thereto contains a signal for that target localization, i.e., a nuclear localization signal.

The prostate cancer antibodies of the invention specifically bind to prostate cancer proteins. By "specifically bind" herein is meant that the antibodies bind to the protein with a  $K_d$  of at least about 0.1 mM, more usually at least about 1  $\mu$ M, preferably at least about 0.1  $\mu$ M or better, and most preferably, 0.01  $\mu$ M or better. Selectivity of binding is also important.

#### **Detection of prostate cancer sequence for diagnostic and therapeutic applications**

In one aspect, the RNA expression levels of genes are determined for different cellular states in the prostate cancer phenotype. After androgen ablation therapy, cells that survive the therapy undergo a period of quiescence followed at sometime later by active cell

division. As explained above, there are a variety of expression patterns characteristic of the prostate cancer genes involved in androgen-independent prostate cancer. Some genes are expressed early in the time course following ablation therapy, then drop off in expression, and then express again with emergence of androgen-independence (hi-lo-hi pattern in 1A).

5 Other genes are expressed early in the time course following ablation therapy, then drop off in expression, and do not express again with emergence of androgen-independence (hi-lo-lo pattern in Table 1A). Still other genes are not expressed early in the time course, but express only with emergence of androgen-independence (lo-lo-hi pattern in Table 1A). Other genes are not expressed early in the time course, but then express as androgen is withdrawn and continue to express with emergence of androgen-independence (lo-hi-hi pattern in Table 1A). Finally, some genes are not expressed early in the time course, but then express as androgen is withdrawn and drop off again with emergence of androgen-independence (lo-hi-lo pattern in Table 1A). Thus, the data suggest that different antigens are expressed in quiescent cells and actively dividing androgen-independent prostate cancer cells.

15 In another aspect, the RNA expression levels of genes are determined for different cellular states in the prostate cancer phenotype. After androgen ablation therapy, cells that survive the therapy undergo a period of quiescence followed at sometime later by active cell division. As explained above, there are a variety of expression patterns characteristic of the prostate cancer genes involved in androgen-independent prostate cancer. Some genes are expressed early in the time course following ablation therapy, then drop off in expression, and then express again with emergence of androgen-independence (hi-lo-lo-hi pattern in Table 2A). Other genes are expressed early in the time course following ablation therapy, then drop off in expression, and do not express again with emergence of androgen-independence (hi-lo-lo-lo and hi-hi-lo-lo pattern in Table 2A). Still other genes are not expressed early in the time course, but express only with emergence of androgen-independence (lo-lo-lo-hi pattern in Table 2A). Other genes are not expressed early in the time course, but then express as androgen is withdrawn and continue to express with emergence of androgen-independence (lo-lo-hi-hi pattern in Table 2A). Finally, some genes are not expressed early in the time course, but then express as androgen is withdrawn and drop off again with emergence of androgen-independence (lo-lo-hi-lo pattern in Table 2A). Thus, the data suggest that different antigens are expressed in quiescent cells (during androgen withdrawal) and actively dividing androgen-independent prostate cancer cells.

Effective therapy to combat androgen-independent prostate cancer requires that the timing of therapy coincide with expression of the target genes. Patients can be monitored for the expression of certain diagnostic antigens that indicate the presence of quiescent cells or which indicate the transition to actively dividing androgen-independent prostate cancer cells.

5 Thus, therapy to combat androgen-independent prostate cancer should begin at some time following androgen ablation therapy, depending on the particular target. Typically the transition from quiescence to actively dividing androgen-independent prostate cancer occurs between 6-24 months following androgen ablation therapy. Thus, preferred time periods for the therapies of the invention are as follows:

10 Expression levels of genes in normal tissue (i.e., not undergoing prostate cancer) and in prostate cancer tissue (and in some cases, for varying severities of prostate cancer that relate to prognosis, as outlined below) or in non-malignant disease are evaluated to provide expression profiles. An expression profile of a particular cell state or point of development is essentially a "fingerprint" of the state. While two states may have a particular gene similarly  
15 expressed, the evaluation of a number of genes simultaneously allows the generation of a gene expression profile that is reflective of the state of the cell. By comparing expression profiles of cells in different states, information regarding which genes are important (including both up- and down-regulation of genes) in each of these states is obtained. Then, diagnosis may be performed or confirmed to determine whether a tissue sample has the gene  
20 expression profile of normal or cancerous tissue. This will provide for molecular diagnosis of related conditions.

"Differential expression," or grammatical equivalents as used herein, refers to qualitative or quantitative differences in the temporal and/or cellular gene expression patterns within and among cells and tissue. Thus, a differentially expressed gene can qualitatively  
25 have its expression altered, including an activation or inactivation, in, e.g., normal versus prostate cancer tissue. Genes may be turned on or turned off in a particular state, relative to another state thus permitting comparison of two or more states. A qualitatively regulated gene will exhibit an expression pattern within a state or cell type which is detectable by standard techniques. Some genes will be expressed in one state or cell type, but not in both.  
30 Alternatively, the difference in expression may be quantitative, e.g., in that expression is increased or decreased; i.e., gene expression is either upregulated, resulting in an increased amount of transcript, or downregulated, resulting in a decreased amount of transcript. The

degree to which expression differs need only be large enough to quantify via standard characterization techniques as outlined below, such as by use of Affymetrix GeneChip™ expression arrays, Lockhart (1996) *Nature Biotechnology* 14:1675-1680, hereby expressly incorporated by reference. Other techniques include, but are not limited to, quantitative reverse transcriptase PCR, northern analysis and RNase protection. As outlined above, preferably the change in expression (i.e., upregulation or downregulation) is at least about 50%, more preferably at least about 100%, more preferably at least about 150%, more preferably at least about 200%, with from 300 to at least 1000% being especially preferred.

Evaluation may be at the gene transcript, or the protein level. The amount of gene expression may be monitored using nucleic acid probes to the DNA or RNA equivalent of the gene transcript, and the quantification of gene expression levels, or, alternatively, the final gene product itself (protein) can be monitored, e.g., with antibodies to the prostate cancer protein and standard immunoassays (ELISAs, etc.) or other techniques, including mass spectroscopy assays, 2D gel electrophoresis assays, etc. Proteins corresponding to prostate cancer genes, i.e., those identified as being important in a prostate cancer or disease phenotype, can be evaluated in a prostate cancer diagnostic test.

In a preferred embodiment, gene expression monitoring is performed simultaneously on a number of genes. Multiple protein expression monitoring can be performed as well. Similarly, these assays may be performed on an individual basis as well.

In this embodiment, the prostate cancer nucleic acid probes are attached to biochips as outlined herein for the detection and quantification of prostate cancer sequences in a particular cell. The assays are further described below in the example. PCR techniques can be used to provide greater sensitivity.

In a preferred embodiment nucleic acids encoding the prostate cancer protein are detected. Although DNA or RNA encoding the prostate cancer protein may be detected, of particular interest are methods wherein an mRNA encoding a prostate cancer protein is detected. Probes to detect mRNA can be a nucleotide/deoxynucleotide probe that is complementary to and hybridizes with the mRNA and includes, but is not limited to, oligonucleotides, cDNA, or RNA. Probes also should contain a detectable label, as defined herein. In one method the mRNA is detected after immobilizing the nucleic acid to be examined on a solid support such as nylon membranes and hybridizing the probe with the sample. Following washing to remove the non-specifically bound probe, the label is

detected. In another method detection of the mRNA is performed in situ (in situ hybridization or ISH). In this method permeabilized cells or tissue samples are contacted with a detectably labeled nucleic acid probe for sufficient time to allow the probe to hybridize with the target mRNA. Following washing to remove the non-specifically bound probe, the label is detected. For example a digoxigenin labeled riboprobe (RNA probe) that is complementary to the mRNA encoding a prostate cancer protein is detected by binding the digoxigenin with an anti-digoxigenin secondary antibody and developed with nitro blue tetrazolium and 5-bromo-4-chloro-3-indoyl phosphate.

In a preferred embodiment, various proteins from the three classes of proteins as described herein (secreted, transmembrane, or intracellular proteins) are used in diagnostic assays. The prostate cancer proteins, antibodies, nucleic acids, modified proteins and cells containing prostate cancer sequences are used in diagnostic assays. Such may evaluate tissues, e.g., immunohistochemistry, or evaluate body fluids, e.g., blood. The detection may be direct of cells, or indirect, e.g., of products from cells. This can be performed on an individual gene or corresponding polypeptide level. In a preferred embodiment, the expression profiles are used, preferably in conjunction with high throughput screening techniques to allow monitoring for expression profile genes and/or corresponding polypeptides.

As described and defined herein, prostate cancer proteins, including intracellular, transmembrane, or secreted proteins, find use as prognostic or diagnostic markers of prostate cancer or other prostate conditions. Detection of these proteins in putative prostate cancer tissue allows for detection, diagnosis, or prognosis of prostate proliferative disorders (malignant and non-malignant) including benign prostate hyperplasia (BPH) and cancer, and prostatitis. Diagnosis may also assist in selecting a therapeutic strategy, e.g., based on expression profiles and/or comparison to archival samples. In one embodiment, antibodies are used to detect prostate cancer proteins, directly or indirectly. A preferred method separates proteins from a sample by electrophoresis on a gel (typically a denaturing and reducing protein gel, but may be another type of gel, including isoelectric focusing gels and the like). Following separation of proteins, the prostate cancer protein is detected, e.g., by immunoblotting with antibodies raised against the prostate cancer protein. Methods of immunoblotting are well known to those of ordinary skill in the art.



In another preferred method, antibodies to the prostate cancer protein find use in in situ imaging techniques, e.g., in histology and/or in immunohistochemistry (e.g., Asai (ed. 1993) Methods in Cell Biology: Antibodies in Cell Biology (vol. 37) Academic Press. In this method cells are contacted with from one to many antibodies to the prostate cancer protein(s).

5 Following washing to remove non-specific antibody binding, the presence of the antibody or antibodies is detected. In one embodiment the antibody is detected by incubating with a secondary antibody that contains a detectable label. In another method the primary antibody to the prostate cancer protein(s) contains a detectable label, e.g., an enzyme marker that can act on a substrate. In another preferred embodiment each one of multiple primary antibodies  
10 contains a distinct and detectable label. This method finds particular use in simultaneous screening for a plurality of prostate cancer proteins. As will be appreciated by one of ordinary skill in the art, many other histological imaging techniques are also provided by the invention.

In a preferred embodiment the label is detected in a fluorometer which has the ability  
15 to detect and distinguish emissions of different wavelengths. In addition, a fluorescence activated cell sorter (FACS) can be used in the method.

In another preferred embodiment, antibodies find use in diagnosing prostate cancer from blood, serum, plasma, stool, and other samples. Such samples, therefore, are useful as samples to be probed or tested for the presence of prostate cancer proteins, which may be  
20 diagnostic of prostate conditions beyond cancer, e.g., BPH. Antibodies can be used to detect a prostate cancer protein by previously described immunoassay techniques including ELISA, immunoblotting (western blotting), immunoprecipitation, BIACORE technology, and the like. Conversely, the presence of antibodies may indicate an immune response against an endogenous prostate cancer protein.

25 In a preferred embodiment, in situ hybridization of labeled prostate cancer nucleic acid probes to tissue arrays is done. For example, arrays of tissue samples, including prostate cancer tissue and/or normal tissue, are made. In situ hybridization (see, e.g., Ausubel, supra) is then performed. When comparing the fingerprints between an individual and a standard, the skilled artisan can make a diagnosis, a prognosis, or a prediction based on the findings. It  
30 is further understood that the genes which indicate the diagnosis may differ from those which indicate the prognosis and molecular profiling of the condition of the cells may lead to distinctions between responsive or refractory conditions or may be predictive of outcomes.

In a preferred embodiment, the prostate cancer proteins, antibodies, nucleic acids, modified proteins, and cells containing prostate cancer sequences are used in prognosis assays. As above, gene expression profiles can be generated that correlate to prostate cancer or other prostate disorders, in terms of useful aspects of clinical condition, pathology, or other information which may be relevant to long term prognosis. Again, this may be done on either a protein or gene level, with the use of genes being preferred. Single or multiple genes may be useful in various combinations. As above, prostate cancer probes may be attached to biochips for the detection and quantification of prostate cancer sequences in a tissue or patient. The assays proceed as outlined above for diagnosis. PCR method may provide more sensitive and accurate quantification.

#### Assays for therapeutic compounds

In a preferred embodiment members of the proteins, nucleic acids, and antibodies as described herein are used in drug screening assays. The prostate cancer proteins, antibodies, nucleic acids, modified proteins, and cells containing prostate cancer sequences are used in drug screening assays or by evaluating the effect of drug candidates on a "gene expression profile" or expression profile of polypeptides. In a preferred embodiment, the expression profiles are used, preferably in conjunction with high throughput screening techniques to allow monitoring for expression profile genes after treatment with a candidate agent (e.g., Zlokarnik, et al. (1998) *Science* 279:84-88; Heid (1996) *Genome Res.* 6:986-94).

In a preferred embodiment, the prostate cancer proteins, antibodies, nucleic acids, modified proteins, and cells containing the native or modified prostate cancer proteins are used in screening assays. That is, the present invention provides novel methods for screening for compositions which modulate the prostate cancer phenotype or an identified physiological function of a prostate cancer protein. As above, this can be done on an individual gene level or by evaluating the effect of drug candidates on a "gene expression profile". In a preferred embodiment, the expression profiles are used, preferably in conjunction with high throughput screening techniques to allow monitoring for expression profile genes after treatment with a candidate agent, see Zlokarnik, supra.

Having identified the differentially expressed genes herein, a variety of assays may be executed. In a preferred embodiment, assays may be run on an individual gene or protein level. That is, having identified a particular gene as up regulated in prostate cancer, test

compounds can be screened for the ability to modulate gene expression or for binding to the prostate cancer protein. "Modulation" thus includes both an increase and a decrease in gene expression. The preferred amount of modulation will depend on the original change of the gene expression in normal versus tissue undergoing prostate cancer, with changes of at least 10%, preferably 50%, more preferably 100-300%, and in some embodiments 300-1000% or greater. Thus, if a gene exhibits a 4-fold increase in prostate cancer tissue compared to normal tissue, a decrease of about four-fold is often desired; similarly, a 10-fold decrease in prostate cancer tissue compared to normal tissue often provides a target value of a 10-fold increase in expression to be induced by the test compound.

The amount of gene expression may be monitored using nucleic acid probes and the quantification of gene expression levels, or, alternatively, the gene product itself can be monitored, e.g., through the use of antibodies to the prostate cancer protein and standard immunoassays. Proteomics and separation techniques may also allow quantification of expression.

In a preferred embodiment, gene expression or protein monitoring of a number of entities, i.e., an expression profile, is monitored simultaneously. Such profiles will typically involve a plurality of those entities described herein.

In this embodiment, the prostate cancer nucleic acid probes are attached to biochips as outlined herein for the detection and quantification of prostate cancer sequences in a particular cell. Alternatively, PCR may be used. Thus, a series, e.g., of microtiter plate, may be used with dispensed primers in desired wells. A PCR reaction can then be performed and analyzed for each well.

Expression monitoring can be performed to identify compounds that modify the expression of one or more prostate cancer-associated sequences, e.g., a polynucleotide sequence set out in Tables 1A-4. Generally, in a preferred embodiment, a test modulator is added to the cells prior to analysis. Moreover, screens are also provided to identify agents that modulate prostate cancer, modulate prostate cancer proteins, bind to a prostate cancer protein, or interfere with the binding of a prostate cancer protein and an antibody or other binding partner.

The term "test compound" or "drug candidate" or "modulator" or grammatical equivalents as used herein describes a molecule, e.g., protein, oligopeptide, small organic molecule, polysaccharide, polynucleotide, etc., to be tested for the capacity to directly or

indirectly alter the prostate cancer phenotype or the expression of a prostate cancer sequence, e.g., a nucleic acid or protein sequence. In preferred embodiments, modulators alter expression profiles, or expression profile nucleic acids or proteins provided herein. In one embodiment, the modulator suppresses a prostate cancer phenotype, e.g., to a normal or non-malignant tissue fingerprint. In another embodiment, a modulator induced a prostate cancer phenotype. Generally, a plurality of assay mixtures are run in parallel with different agent concentrations to obtain a differential response to the various concentrations. Typically, one of these concentrations serves as a negative control, i.e., at zero concentration or below the level of detection.

Drug candidates encompass numerous chemical classes, though typically they are organic molecules, preferably small organic compounds having a molecular weight of more than 100 and less than about 2,500 daltons. Preferred small molecules are less than 2000, or less than 1500, or less than 1000, or less than 500 D. Candidate agents comprise functional groups necessary for structural interaction with proteins, particularly hydrogen bonding, and typically include at least an amine, carbonyl, hydroxyl or carboxyl group, preferably at least two of the functional chemical groups. The candidate agents often comprise cyclical carbon or heterocyclic structures and/or aromatic or polyaromatic structures substituted with one or more of the above functional groups. Candidate agents are also found among biomolecules including peptides, saccharides, fatty acids, steroids, purines, pyrimidines, derivatives, structural analogs, or combinations thereof. Particularly preferred are peptides.

In one aspect, a modulator will neutralize the effect of a prostate cancer protein. By "neutralize" is meant that activity of a protein is inhibited or blocked and the consequent effect on the cell.

In certain embodiments, combinatorial libraries of potential modulators will be screened for an ability to bind to a prostate cancer polypeptide or to modulate activity. Conventionally, new chemical entities with useful properties are generated by identifying a chemical compound (called a "lead compound") with some desirable property or activity, e.g., inhibiting activity, creating variants of the lead compound, and evaluating the property and activity of those variant compounds. Often, high throughput screening (HTS) methods are employed for such an analysis.

In one preferred embodiment, high throughput screening methods involve providing a library containing a large number of potential therapeutic compounds (candidate

compounds). Such "combinatorial chemical libraries" are then screened in one or more assays to identify those library members (particular chemical species or subclasses) that display a desired characteristic activity. The compounds thus identified can serve as conventional "lead compounds" or can themselves be used as potential or actual therapeutics.

5 A combinatorial chemical library is a collection of diverse chemical compounds generated by either chemical synthesis or biological synthesis by combining a number of chemical "building blocks" such as reagents. For example, a linear combinatorial chemical library, such as a polypeptide (e.g., mutein) library, is formed by combining a set of chemical building blocks called amino acids in most every possible way for a given compound length  
10 (i.e., the number of amino acids in a polypeptide compound). Millions of chemical compounds can be synthesized through such combinatorial mixing of chemical building blocks. Gallop, et al. (1994) J. Med. Chem. 37:1233-1251.

Preparation and screening of combinatorial chemical libraries is well known to those of skill in the art. Such combinatorial chemical libraries include, but are not limited to,  
15 peptide libraries (see, e.g., U.S. Patent No. 5,010,175, Furka (1991) Pept. Prot. Res. 37:487-493, Houghton, et al. (1991) Nature, 354:84-88), peptoids (PCT Publication No WO 91/19735), encoded peptides (PCT Publication WO 93/20242), random bio-oligomers (PCT Publication WO 92/00091), benzodiazepines (U.S. Pat. No. 5,288,514), diversomers such as hydantoins, benzodiazepines and dipeptides (Hobbs, et al. (1993) Proc. Nat. Acad. Sci. USA  
20 90:6909-6913), vinylogous polypeptides (Hagihara, et al. (1992) J. Amer. Chem. Soc. 114:6568-xxx), nonpeptidic peptidomimetics with a Beta-D-Glucose scaffolding (Hirschmann, et al. (1992) J. Amer. Chem. Soc. 114:9217-9218), analogous organic syntheses of small compound libraries (Chen, et al. (1994) J. Amer. Chem. Soc. 116:2661-xxx), oligocarbamates (Cho, et al. (1993) Science 261:1303-1305), and/or peptidyl  
25 phosphonates (Campbell, et al. (1994) J. Org. Chem. 59:658-xxx). See, generally, Gordon, et al. (1994) J. Med. Chem. 37:1385-1401), nucleic acid libraries (see, e.g., Stratagene, Corp.), peptide nucleic acid libraries (see, e.g., U.S. Patent 5,539,083), antibody libraries (see, e.g., Vaughn, et al. (1996) Nature Biotechnology 14:309-314, and PCT/US96/10287), carbohydrate libraries (see, e.g., Liang, et al. (1996) Science 274:1520-1522, and U.S. Patent  
30 No. 5,593,853), and small organic molecule libraries (see, e.g., benzodiazepines, Baum (1993) C&EN, Jan 18, page 33; isoprenoids, U.S. Patent No. 5,569,588; thiazolidinones and metathiazanones, U.S. Patent No. 5,549,974; pyrrolidines, U.S. Patent Nos. 5,525,735 and

5,519,134; morpholino compounds, U.S. Patent No. 5,506,337; benzodiazepines, U.S. Patent No. 5,288,514; and the like).

Devices for the preparation of combinatorial libraries are commercially available (see, e.g., 357 MPS, 390 MPS, Advanced Chem Tech, Louisville KY, Symphony, Rainin, Woburn, MA, 433A Applied Biosystems, Foster City, CA, 9050 Plus, Millipore, Bedford, MA).

A number of well known robotic systems have also been developed for solution phase chemistries. These systems include automated workstations like the automated synthesis apparatus developed by Takeda Chemical Industries, LTD. (Osaka, Japan) and many robotic systems utilizing robotic arms (Zymate II, Zymark Corporation, Hopkinton, Mass.; Orca, Hewlett-Packard, Palo Alto, Calif.), which mimic the manual synthetic operations performed by a chemist. Many of the above devices are suitable for use with the present invention. The nature and implementation of modifications to these devices (if any) so that they can operate as discussed herein will be apparent to persons skilled in the relevant art. In addition, numerous combinatorial libraries are themselves commercially available (see, e.g., ComGenex, Princeton, N.J., Asinex, Moscow, Ru, Tripos, Inc., St. Louis, MO, ChemStar, Ltd, Moscow, RU, 3D Pharmaceuticals, Exton, PA, Martek Biosciences, Columbia, MD, etc.).

The assays to identify modulators are amenable to high throughput screening.

Preferred assays thus detect enhancement or inhibition of prostate cancer gene transcription, inhibition or enhancement of polypeptide expression, and inhibition or enhancement of polypeptide activity.

High throughput assays for the presence, absence, quantification, or other properties of particular nucleic acids or protein products are well known to those of skill in the art.

Similarly, binding assays and reporter gene assays are similarly well known. Thus, e.g., U.S. Patent No. 5,559,410 discloses high throughput screening methods for proteins, U.S. Patent No. 5,585,639 discloses high throughput screening methods for nucleic acid binding (i.e., in arrays), while U.S. Patent Nos. 5,576,220 and 5,541,061 disclose high throughput methods of screening for ligand/antibody binding.

In addition, high throughput screening systems are commercially available (see, e.g., Zymark Corp., Hopkinton, MA; Air Technical Industries, Mentor, OH; Beckman Instruments, Inc. Fullerton, CA; Precision Systems, Inc., Natick, MA, etc.). These systems

typically automate entire procedures, including sample and reagent pipetting, liquid dispensing, timed incubations, and final readings of the microplate in detector(s) appropriate for the assay. These configurable systems provide high throughput and rapid start up as well as a high degree of flexibility and customization. The manufacturers of such systems provide  
5 detailed protocols for various high throughput systems. Thus, e.g., Zymark Corp. provides technical bulletins describing screening systems for detecting the modulation of gene transcription, ligand binding, and the like.

In one embodiment, modulators are proteins, often naturally occurring proteins or fragments of naturally occurring proteins. Thus, e.g., cellular extracts containing proteins, or  
10 random or directed digests of proteinaceous cellular extracts, may be used. In this way libraries of proteins may be made for screening in the methods of the invention. Particularly preferred in this embodiment are libraries of bacterial, fungal, viral, and mammalian proteins, with the latter being preferred, and human proteins being especially preferred. Particularly useful test compound will be directed to the class of proteins to which the target belongs, e.g.,  
15 substrates for enzymes or ligands and receptors.

In a preferred embodiment, modulators are peptides of from about 5 to about 30 amino acids, with from about 5 to about 20 amino acids being preferred, and from about 7 to about 15 being particularly preferred. The peptides may be digests of naturally occurring proteins as is outlined above, random peptides, or "biased" random peptides. By  
20 "randomized" or grammatical equivalents herein is meant that each nucleic acid and peptide consists of essentially random nucleotides and amino acids, respectively. Since generally these random peptides (or nucleic acids, discussed below) are chemically synthesized, they may typically incorporate any nucleotide or amino acid at any position. The synthetic process can be designed to generate randomized proteins or nucleic acids, to allow the  
25 formation of all or most of the possible combinations over the length of the sequence, thus forming a library of randomized candidate bioactive proteinaceous agents.

In one embodiment, the library is fully randomized, with no sequence preferences or constants at any position. In a preferred embodiment, the library is biased. That is, some positions within the sequence are either held constant, or are selected from a limited number  
30 of possibilities. For example, in a preferred embodiment, the nucleotides or amino acid residues are randomized within a defined class, e.g., of hydrophobic amino acids, hydrophilic residues, sterically biased (either small or large) residues, towards the creation of nucleic acid

binding domains, the creation of cysteines, for cross-linking, prolines for SH-3 domains, serines, threonines, tyrosines, or histidines for phosphorylation sites, etc., or to purines, etc.

Modulators of prostate cancer can also be nucleic acids, as defined above.

As described above generally for proteins, nucleic acid modulating agents may be  
5 naturally occurring nucleic acids, random nucleic acids, or "biased" random nucleic acids. For example, digests of prokaryotic or eukaryotic genomes may be used as is outlined above for proteins.

In a preferred embodiment, the candidate compounds are organic chemical moieties, a wide variety of which are available in the literature.

10 After the candidate agent has been added and the cells allowed to incubate for some period of time, the sample containing a target sequence to be analyzed is added to the biochip. If required, the target sequence is prepared using known techniques. For example, the sample may be treated to lyse the cells, using known lysis buffers, electroporation, etc., with purification and/or amplification such as PCR performed as appropriate. For example,  
15 an in vitro transcription with labels covalently attached to the nucleotides is performed. Generally, the nucleic acids are labeled with biotin-FITC or PE, or with cy3 or cy5.

In a preferred embodiment, the target sequence is labeled with, e.g., a fluorescent, a chemiluminescent, a chemical, or a radioactive signal, to provide a means of detecting the target sequence's specific binding to a probe. The label also can be an enzyme, such as,  
20 alkaline phosphatase or horseradish peroxidase, which when provided with an appropriate substrate produces a product that can be detected. Alternatively, the label can be a labeled compound or small molecule, such as an enzyme inhibitor, that binds but is not catalyzed or altered by the enzyme. The label also can be a moiety or compound, such as, an epitope tag or biotin which specifically binds to streptavidin. For the example of biotin, the streptavidin  
25 is labeled as described above, thereby, providing a detectable signal for the bound target sequence. Unbound labeled streptavidin is typically removed prior to analysis.

As will be appreciated by those in the art, these assays can be direct hybridization assays or can comprise "sandwich assays", which include the use of multiple probes, as is generally outlined in U.S. Patent Nos. 5,681,702, 5,597,909, 5,545,730, 5,594,117,  
30 5,591,584, 5,571,670, 5,580,731, 5,571,670, 5,591,584, 5,624,802, 5,635,352, 5,594,118, 5,359,100, 5,124,246, and 5,681,697, each of which is hereby incorporated by reference. In this embodiment, in general, the target nucleic acid is prepared as outlined above, and then



added to the biochip comprising a plurality of nucleic acid probes, under conditions that allow the formation of a hybridization complex.

A variety of hybridization conditions may be used in the present invention, including high, moderate, and low stringency conditions as outlined above. The assays are generally run under stringency conditions which allows formation of the label probe hybridization complex only in the presence of target. Stringency can be controlled by altering a step parameter that is a thermodynamic variable, including, but not limited to, temperature, formamide concentration, salt concentration, chaotropic salt concentration pH, organic solvent concentration, etc.

These parameters may also be used to control non-specific binding, as is generally outlined in U.S. Patent No. 5,681,697. Thus it may be desirable to perform certain steps at higher stringency conditions to reduce non-specific binding.

The reactions outlined herein may be accomplished in a variety of ways. Components of the reaction may be added simultaneously, or sequentially, in different orders, with preferred embodiments outlined below. In addition, the reaction may include a variety of other reagents. These include salts, buffers, neutral proteins, e.g., albumin, detergents, etc., which may be used to facilitate optimal hybridization and detection, and/or reduce non-specific or background interactions. Reagents that otherwise improve the efficiency of the assay, such as protease inhibitors, nuclease inhibitors, anti-microbial agents, etc., may also be used as appropriate, depending on the sample preparation methods and purity of the target.

The assay data are analyzed to determine the expression levels, and changes in expression levels as between states, of individual genes, forming a gene expression profile.

Screens are performed to identify modulators of the prostate cancer or related phenotype. In one embodiment, screening is performed to identify modulators that can induce or suppress a particular expression profile, thus preferably generating the associated phenotype. In another embodiment, e.g., for diagnostic applications, having identified differentially expressed genes important in a particular state, screens can be performed to identify modulators that alter expression of individual genes. In an another embodiment, screening is performed to identify modulators that alter a biological function of the expression product of a differentially expressed gene. Again, having identified the importance of a gene in a particular state, screens are performed to identify agents that bind and/or modulate the biological activity of the gene product.

In addition screens can be done for genes that are induced in response to a candidate agent. After identifying a modulator based upon its ability to suppress a prostate cancer expression pattern leading to a normal expression pattern, or to modulate a single prostate cancer gene expression profile so as to mimic the expression of the gene from normal tissue, a screen as described above can be performed to identify genes that are specifically modulated in response to the agent. Comparing expression profiles between normal tissue and agent treated prostate cancer tissue reveals genes that are not expressed in normal tissue or prostate cancer tissue, but are expressed in agent treated tissue. These agent-specific sequences can be identified and used by methods described herein for prostate cancer genes or proteins. In particular these sequences and the proteins they encode find use in marking or identifying agent treated cells. In addition, antibodies can be raised against the agent induced proteins and used to target novel therapeutics to the treated prostate cancer tissue sample.

Thus, in one embodiment, a test compound is administered to a population of prostate cancer cells, that have an associated prostate cancer expression profile. By "administration" or "contacting" herein is meant that the candidate agent is added to the cells in such a manner as to allow the agent to act upon the cell, whether by uptake and intracellular action, or by action at the cell surface. In some embodiments, nucleic acid encoding a proteinaceous candidate agent (e.g., a peptide) may be put into a viral construct such as an adenoviral or retroviral construct, and added to the cell, such that expression of the peptide agent is accomplished, e.g., PCT US97/01019. Regulatable gene therapy systems can also be used.

Once the test compound has been administered to the cells, the cells can be washed if desired and are allowed to incubate under preferably physiological conditions for some period of time. The cells are then harvested and a new gene expression profile is generated, as outlined herein.

Thus, e.g., prostate cancer or non-malignant tissue may be screened for agents that modulate, e.g., induce or suppress the prostate cancer or related phenotype. A change in at least one gene, preferably many, of the expression profile indicates that the agent has an effect on prostate cancer activity. By defining such a signature for the prostate cancer phenotype, screens for new drugs that alter the phenotype can be devised. With this approach, the drug target need not be known and need not be represented in the original expression screening platform, nor does the level of transcript for the target protein need to change.

In a preferred embodiment, as outlined above, screens may be done on individual genes and gene products (proteins). That is, having identified a particular differentially expressed gene as important in a particular state, screening of modulators of either the expression of the gene or the gene product itself can be done. The gene products of differentially expressed genes are sometimes referred to herein as "prostate cancer proteins" or a "prostate cancer modulatory protein". The prostate cancer modulatory protein may be a fragment, or alternatively, be the full length protein to the fragment encoded by the nucleic acids of the Tables 1A-4. Preferably, the prostate cancer modulatory protein is a fragment. In a preferred embodiment, the prostate cancer amino acid sequence which is used to determine sequence identity or similarity is encoded by a nucleic acid of Tables 1A-4. In another embodiment, the sequences are naturally occurring allelic variants of a protein encoded by a nucleic acid of Tables 1A-4. In another embodiment, the sequences are sequence variants as further described herein.

Preferably, the prostate cancer modulatory protein is a fragment of approximately 14 to 24 amino acids long. More preferably the fragment is a soluble fragment. Preferably, the fragment includes a non-transmembrane region. In a preferred embodiment, the fragment has an N-terminal Cys to aid in solubility. In one embodiment, the C-terminus of the fragment is kept as a free acid and the N-terminus is a free amine to aid in coupling, i.e., to cysteine.

In one embodiment the prostate cancer proteins are conjugated to an immunogenic agent as discussed herein. In one embodiment the prostate cancer protein is conjugated to BSA.

Measurements of prostate cancer polypeptide activity, or of prostate cancer or the prostate cancer phenotype can be performed using a variety of assays. For example, the effects of the test compounds upon the function of the prostate cancer polypeptides can be measured by examining parameters described above. A suitable physiological change that affects activity can be used to assess the influence of a test compound on the polypeptides of this invention. When the functional consequences are determined using intact cells or animals, one can also measure a variety of effects such as, in the case of prostate cancer associated with tumors, tumor growth, tumor metastasis, neovascularization, hormone release, transcriptional changes to both known and uncharacterized genetic markers (e.g., northern blots), changes in cell metabolism such as cell growth or pH changes, and changes

in intracellular second messengers such as cGMP. In the assays of the invention, a mammalian prostate cancer polypeptide is typically used, e.g., mouse, preferably human.

Assays to identify compounds with modulating activity can be performed in vitro. For example, a prostate cancer polypeptide is first contacted with a potential modulator and incubated for a suitable amount of time, e.g., from 0.5 to 48 hours. In one embodiment, the prostate cancer polypeptide levels are determined in vitro by measuring the level of protein or mRNA. The level of protein is measured using immunoassays such as western blotting, ELISA, and the like with an antibody that selectively binds to the prostate cancer polypeptide or a fragment thereof. For measurement of mRNA, amplification, e.g., using PCR, LCR, or hybridization assays, e.g., northern hybridization, RNase protection, dot blotting, are preferred. The level of protein or mRNA is detected using directly or indirectly labeled detection agents, e.g., fluorescently or radioactively labeled nucleic acids, radioactively or enzymatically labeled antibodies, and the like, as described herein.

Alternatively, a reporter gene system can be devised using the prostate cancer protein promoter operably linked to a reporter gene such as luciferase, green fluorescent protein, CAT, or  $\beta$ -gal. The reporter construct is typically transfected into a cell. After treatment with a potential modulator, the amount of reporter gene transcription, translation, or activity is measured according to standard techniques known to those of skill in the art.

In a preferred embodiment, as outlined above, screens may be done on individual genes and gene products (proteins). That is, having identified a particular differentially expressed gene as important in a particular state, screening of modulators of the expression of the gene or the gene product itself can be done. The gene products of differentially expressed genes are sometimes referred to herein as "prostate cancer proteins." The prostate cancer protein may be a fragment, or alternatively, be the full length protein corresponding to a fragment shown herein.

In one embodiment, screening for modulators of expression of specific genes is performed. Typically, the expression of only one or a few genes are evaluated. In another embodiment, screens are designed to first find compounds that bind to differentially expressed proteins. These compounds are then evaluated for the ability to modulate differentially expressed activity. Moreover, once initial candidate compounds are identified, variants can be further screened to better evaluate structure activity relationships.

In a preferred embodiment, binding assays are done. In general, purified or isolated gene product is used; that is, the gene products of one or more differentially expressed nucleic acids are made. For example, antibodies are generated to the protein gene products, and standard immunoassays are run to determine the amount of protein present.

- 5 Alternatively, cells comprising the prostate cancer proteins can be used in the assays.

Thus, in a preferred embodiment, the methods comprise combining a prostate cancer protein and a candidate compound, and determining the binding of the compound to the prostate cancer protein. Preferred embodiments utilize the human prostate cancer protein, although other mammalian proteins may also be used, e.g., for the development of animal  
10 models of human disease. In some embodiments, as outlined herein, variant or derivative prostate cancer proteins may be used.

Generally, in a preferred embodiment of the methods herein, the prostate cancer protein or the candidate agent is non-diffusably bound to an insoluble support having isolated sample receiving areas (e.g., a microtiter plate, an array, etc.). The insoluble supports may be  
15 made of a composition to which the compositions can be bound, is readily separated from soluble material, and is otherwise compatible with the overall method of screening. The surface of such supports may be solid or porous and of a convenient shape. Examples of suitable insoluble supports include microtiter plates, arrays, membranes, and beads. These are typically made of glass, plastic (e.g., polystyrene), polysaccharides, nylon or  
20 nitrocellulose, teflon™, etc. Microtiter plates and arrays are especially convenient because a large number of assays can be carried out simultaneously, using small amounts of reagents and samples. The particular manner of binding of the composition should be compatible with the reagents and overall methods of the invention, maintain the activity of the composition, and be nondiffusable. Preferred methods of binding include the use of antibodies (which do  
25 not sterically block either the ligand binding site or activation sequence when the protein is bound to the support), direct binding to "sticky" or ionic supports, chemical crosslinking, the synthesis of the protein or agent on the surface, etc. Following binding of the protein or agent, excess unbound material is removed by washing. The sample receiving areas may then be blocked through incubation with bovine serum albumin (BSA), casein, or other  
30 innocuous protein or other moiety.

In a preferred embodiment, the prostate cancer protein is bound to the support, and a test compound is added to the assay. Alternatively, the candidate agent is bound to the

support and the prostate cancer protein is added. Novel binding agents include specific antibodies, non-natural binding agents identified in screens of chemical libraries, peptide analogs, etc. Of particular interest are screening assays for agents that have a low toxicity for human cells. A wide variety of assays may be used for this purpose, including labeled in vitro protein-protein binding assays, electrophoretic mobility shift assays, immunoassays for protein binding, functional assays (phosphorylation assays, etc.) and the like.

The determination of the binding of the test modulating compound to the prostate cancer protein may be done in a number of ways. In a preferred embodiment, the compound is labeled, and binding determined directly, e.g., by attaching all or a portion of the prostate cancer protein to a solid support, adding a labeled candidate agent (e.g., a fluorescent label), washing off excess reagent, and determining whether the label is present on the solid support. Various blocking and washing steps may be utilized as appropriate.

In some embodiments, only one of the components is labeled, e.g., the proteins (or proteinaceous candidate compounds) can be labeled. Alternatively, more than one component can be labeled with different labels, e.g.,  $^{125}\text{I}$  for the proteins and a fluorophor for the compound. Proximity reagents, e.g., quenching or energy transfer reagents are also useful.

In one embodiment, the binding of the test compound is determined by competitive binding assay. The competitor is a binding moiety known to bind to the target molecule (i.e., a prostate cancer protein), such as an antibody, peptide, binding partner, ligand, etc. Under certain circumstances, there may be competitive binding between the compound and the binding moiety, with the binding moiety displacing the compound. In one embodiment, the test compound is labeled. Either the compound, or the competitor, or both, is added first to the protein for a time sufficient to allow binding, if present. Incubations may be performed at a temperature which facilitates optimal activity, typically between 4 and 40° C. Incubation periods are typically optimized, e.g., to facilitate rapid high throughput screening. Typically between 0.1 and 1 hour will be sufficient. Excess reagent is generally removed or washed away. The second component is then added, and the presence or absence of the labeled component is followed, to indicate binding.

In a preferred embodiment, the competitor is added first, followed by the test compound. Displacement of the competitor is an indication that the test compound is binding to the prostate cancer protein and thus is capable of binding to, and potentially modulating,

the activity of the prostate cancer protein. In this embodiment, either component can be labeled. Thus, e.g., if the competitor is labeled, the presence of label in the wash solution indicates displacement by the agent. Alternatively, if the test compound is labeled, the presence of the label on the support indicates displacement.

5 In an alternative embodiment, the test compound is added first, with incubation and washing, followed by the competitor. The absence of binding by the competitor may indicate that the test compound is bound to the prostate cancer protein with a higher affinity. Thus, if the test compound is labeled, the presence of the label on the support, coupled with a lack of competitor binding, may indicate that the test compound is capable of binding to the prostate cancer protein.

10 In a preferred embodiment, the methods comprise differential screening to identify agents that are capable of modulating the activity of the prostate cancer proteins. In this embodiment, the methods comprise combining a prostate cancer protein and a competitor in a first sample. A second sample comprises a test compound, a prostate cancer protein, and a competitor. The binding of the competitor is determined for both samples, and a change, or difference in binding between the two samples indicates the presence of an agent capable of binding to the prostate cancer protein and potentially modulating its activity. That is, if the binding of the competitor is different in the second sample relative to the first sample, the agent is capable of binding to the prostate cancer protein.

20 Alternatively, differential screening is used to identify drug candidates that bind to the native prostate cancer protein, but cannot bind to modified prostate cancer proteins. The structure of the prostate cancer protein may be modeled, and used in rational drug design to synthesize agents that interact with that site. Drug candidates that affect the activity of a prostate cancer protein are also identified by screening drugs for the ability to either enhance or reduce the activity of the protein.

25 Positive controls and negative controls may be used in the assays. Preferably control and test samples are performed in at least triplicate to obtain statistically significant results. Incubation of samples is for a time sufficient for the binding of the agent to the protein. Following incubation, samples are washed free of non-specifically bound material and the amount of bound, generally labeled agent determined. For example, where a radiolabel is employed, the samples may be counted in a scintillation counter to determine the amount of bound compound.

A variety of other reagents may be included in the screening assays. These include reagents like salts, neutral proteins, e.g., albumin, detergents, etc., which may be used to facilitate optimal protein-protein binding and/or reduce non-specific or background interactions. Also reagents that otherwise improve the efficiency of the assay, such as

5 protease inhibitors, nuclease inhibitors, anti-microbial agents, etc., may be used. The mixture of components may be added in an order that provides for the requisite binding.

In a preferred embodiment, the invention provides methods for screening for a compound capable of modulating the activity of a prostate cancer protein. The methods comprise adding a test compound, as defined above, to a cell comprising prostate cancer

10 proteins. Preferred cell types include almost any cell. The cells contain a recombinant nucleic acid that encodes a prostate cancer protein. In a preferred embodiment, a library of candidate agents are tested on a plurality of cells.

In one aspect, the assays are evaluated in the presence or absence or previous or subsequent exposure of physiological signals, e.g., hormones, antibodies, peptides, antigens,

15 cytokines, growth factors, action potentials, pharmacological agents including chemotherapeutics, radiation, carcinogenics, or other cells (e.g., cell-cell contacts). In another example, the determinations are determined at different stages of the cell cycle process.

In this way, compounds that modulate prostate cancer agents are identified.

20 Compounds with pharmacological activity are able to enhance or interfere with the activity of the prostate cancer protein. Once identified, similar structures are evaluated to identify critical structural feature of the compound.

In one embodiment, a method of inhibiting prostate cancer cell division is provided. The method comprises administration of a prostate cancer inhibitor. In another embodiment,

25 a method of inhibiting prostate cancer or other prostate proliferative condition is provided. The method comprises administration of a prostate cancer inhibitor. In a further embodiment, methods of treating cells or individuals with prostate cancer are provided. The method comprises administration of a prostate cancer inhibitor.

In one embodiment, a prostate cancer inhibitor is an antibody as discussed above. In

30 another embodiment, the prostate cancer inhibitor is an antisense molecule.

A variety of cell growth, proliferation, and metastasis assays are known to those of skill in the art, as described below.



#### Soft agar growth or colony formation in suspension

Normal cells require a solid substrate to attach and grow. When the cells are transformed, they lose this phenotype and grow detached from the substrate. For example, transformed cells can grow in stirred suspension culture or suspended in semi-solid media, such as semi-solid or soft agar. The transformed cells, when transfected with tumor suppressor genes, regenerate normal phenotype and require a solid substrate to attach and grow. Soft agar growth or colony formation in suspension assays can be used to identify modulators of prostate cancer sequences, which when expressed in host cells, inhibit abnormal cellular proliferation and transformation. A therapeutic compound would reduce or eliminate the host cells' ability to grow in stirred suspension culture or suspended in semi-solid media, such as semi-solid or soft.

Techniques for soft agar growth or colony formation in suspension assays are described in Freshney (1994) Culture of Animal Cells a Manual of Basic Technique 3d ed. Wiley-Liss, herein incorporated by reference. See also, the methods section of Garkavtsev, et al. (1996), supra, herein incorporated by reference.

#### Contact inhibition and density limitation of growth

Normal cells typically grow in a flat and organized pattern in a petri dish until they touch other cells. When the cells touch one another, they are contact inhibited and stop growing. When cells are transformed, however, the cells are not contact inhibited and continue to grow to high densities in disorganized foci. Thus, the transformed cells grow to a higher saturation density than normal cells. This can be detected morphologically by the formation of a disoriented monolayer of cells or rounded cells in foci within the regular pattern of normal surrounding cells. Alternatively, labeling index with ( $^3\text{H}$ )-thymidine at saturation density can be used to measure density limitation of growth. See Freshney (1994), supra. The transformed cells, when transfected with tumor suppressor genes, regenerate a normal phenotype and become contact inhibited and would grow to a lower density.

In this assay, labeling index with ( $^3\text{H}$ )-thymidine at saturation density is a preferred method of measuring density limitation of growth. Transformed host cells are transfected with a prostate cancer-associated sequence and are grown for 24 hours at saturation density in non-limiting medium conditions. The percentage of cells labeling with ( $^3\text{H}$ )-thymidine is determined autoradiographically. See, Freshney (1994), supra.

#### Growth factor or serum dependence

Transformed cells have a lower serum dependence than their normal counterparts (see, e.g., Temin (1966) J. Natl. Cancer Inst. 37:167-175; Eagle, et al. (1970) J. Exp. Med. 131:836-879); Freshney, supra. This is in part due to release of various growth factors by the transformed cells. Growth factor or serum dependence of transformed host cells can be compared with that of control.

#### Tumor specific markers levels

Tumor cells release an increased amount of certain factors (hereinafter "tumor specific markers") than their normal counterparts. For example, plasminogen activator (PA) is released from human glioma at a higher level than from normal brain cells (see, e.g., Gullino, "Angiogenesis, tumor vascularization, and potential interference with tumor growth" pp. 178-184 in Mihich (ed. 1985) Biological Responses in Cancer Plenum. Similarly, Tumor angiogenesis factor (TAF) is released at a higher level in tumor cells than their normal counterparts. See, e.g., Folkman (1992) Angiogenesis and Cancer, Sem. Cancer Biol.

Various techniques which measure the release of these factors are described in Freshney (1994), supra. Also, see, Unkless, et al. (1974) J. Biol. Chem. 249:4295-4305; Strickland and Beers (1976) J. Biol. Chem. 251:5694-5702; Whur, et al. (1980) Br. J. Cancer 42:305-312; Gullino, "Angiogenesis, tumor vascularization, and potential interference with tumor growth" pp. 178-184 in Mihich (ed. 1985) Biological Responses in Cancer Plenum; and Freshney (1985) Anticancer Res. 5:111-130.

#### Invasiveness into Matrigel

The degree of invasiveness into Matrigel or some other extracellular matrix constituent can be used as an assay to identify compounds that modulate prostate cancer-associated sequences. Tumor cells exhibit a good correlation between malignancy and invasiveness of cells into Matrigel or some other extracellular matrix constituent. In this assay, tumorigenic cells are typically used as host cells. Expression of a tumor suppressor gene in these host cells would decrease invasiveness of the host cells.

Techniques described in Freshney (1994), supra, can be used. Briefly, the level of invasion of host cells can be measured by using filters coated with Matrigel or some other extracellular matrix constituent. Penetration into the gel, or through to the distal side of the filter, is rated as invasiveness, and rated histologically by number of cells and distance moved, or by prelabeling the cells with <sup>125</sup>I and counting the radioactivity on the distal side of the filter or bottom of the dish. See, e.g., Freshney (1984), supra.

### Tumor growth in vivo

Effects of prostate cancer-associated sequences on cell growth can be tested in transgenic or immune-suppressed mice. Knock-out transgenic mice can be made, in which the prostate cancer gene is disrupted or in which a prostate cancer gene is inserted. Knock-out transgenic mice can be made by insertion of a marker gene or other heterologous gene into the endogenous prostate cancer gene site in the mouse genome via homologous recombination. Such mice can also be made by substituting the endogenous prostate cancer gene with a mutated version of the prostate cancer gene, or by mutating the endogenous prostate cancer gene, e.g., by exposure to carcinogens.

A DNA construct is introduced into the nuclei of embryonic stem cells. Cells containing the newly engineered genetic lesion are injected into a host mouse embryo, which is re-implanted into a recipient female. Some of these embryos develop into chimeric mice that possess germ cells partially derived from the mutant cell line. Therefore, by breeding the chimeric mice it is possible to obtain a new line of mice containing the introduced genetic lesion (see, e.g., Capecchi, et al. (1989) Science 244:1288-1292). Chimeric targeted mice can be derived according to Hogan, et al. (1988) Manipulating the Mouse Embryo: A Laboratory Manual CSH Press; and Robertson (ed. 1987) Teratocarcinomas and Embryonic Stem Cells: A Practical Approach IRL Press, Washington, D.C.

Alternatively, various immune-suppressed or immune-deficient host animals can be used. For example, genetically athymic "nude" mouse (see, e.g., Giovanella, et al. (1974) J. Natl. Cancer Inst. 52:921-930), a SCID mouse, a thymectomized mouse, or an irradiated mouse (see, e.g., Bradley, et al. (1978) Br. J. Cancer 38:263-272; Selby, et al. (1980) Br. J. Cancer 41:52-61) can be used as a host. Transplantable tumor cells (typically about  $10^6$  cells) injected into isogenic hosts will produce invasive tumors in a high proportions of cases, while normal cells of similar origin will not. In hosts which developed invasive tumors, cells expressing a prostate cancer-associated sequences are injected subcutaneously. After a suitable length of time, preferably 4-8 weeks, tumor growth is measured (e.g., by volume or by its two largest dimensions) and compared to the control. Tumors that have statistically significant reduction (using, e.g., Student's T test) are said to have inhibited growth.

Polynucleotide modulators of prostate cancer

Antisense and RNAi Polynucleotides

In certain embodiments, the activity of a prostate cancer-associated protein is down-regulated, or entirely inhibited, by the use of antisense polynucleotide, i.e., a nucleic acid complementary to, and which can preferably hybridize specifically to, a coding mRNA nucleic acid sequence, e.g., a prostate cancer protein mRNA, or a subsequence thereof.

- 5 Binding of the antisense polynucleotide to the mRNA reduces the translation and/or stability of the mRNA.

In the context of this invention, antisense polynucleotides can comprise naturally-occurring nucleotides, or synthetic species formed from naturally-occurring subunits or their close homologs. Antisense polynucleotides may also have altered sugar moieties or inter-sugar linkages. Exemplary among these are the phosphorothioate and other sulfur containing species which are known for use in the art. Analogs are comprehended by this invention so long as they function effectively to hybridize with the prostate cancer protein mRNA. See, e.g., Isis Pharmaceuticals, Carlsbad, CA; Sequitor, Inc., Natick, MA.

10 Such antisense polynucleotides can readily be synthesized using recombinant means, or can be synthesized in vitro. Equipment for such synthesis is sold by several vendors, including Applied Biosystems. The preparation of other oligonucleotides such as phosphorothioates and alkylated derivatives is also well known to those of skill in the art.

Antisense molecules as used herein include antisense or sense oligonucleotides. Sense oligonucleotides can, e.g., be employed to block transcription by binding to the antisense strand. The antisense and sense oligonucleotide comprise a single-stranded nucleic acid sequence (either RNA or DNA) capable of binding to target mRNA (sense) or DNA (antisense) sequences for prostate cancer molecules. A preferred antisense molecule is for a prostate cancer sequences in Tables 1A-4, or for a ligand or activator thereof. Antisense or sense oligonucleotides, according to the present invention, comprise a fragment generally at least about 14 nucleotides, preferably from about 14 to 30 nucleotides. The ability to derive an antisense or a sense oligonucleotide, based upon a cDNA sequence encoding a given protein is described in, e.g., Stein and Cohen (1988) Cancer Res. 48:2659-2668; and van der Krol, et al. (1988) BioTechniques 6:958-976.

RNA interference is a mechanism to suppress gene expression in a sequence specific manner. See, e.g., Brumelkamp, et al. (2002) Sciencexpress (21March2002); Sharp (1999) Genes Dev. 13:139-141; and Cathew (2001) Curr. Op. Cell Biol. 13:244-248. In mammalian cells, short, e.g., 21 nt, double stranded small interfering RNAs (siRNA) have been shown to

be effective at inducing an RNAi response. See, e.g., Elbashir, et al. (2001) Nature 411:494-498. The mechanism may be used to downregulate expression levels of identified genes, e.g., treatment of or validation of relevance to disease.

## 5 Ribozymes

In addition to antisense polynucleotides, ribozymes can be used to target and inhibit transcription of prostate cancer-associated nucleotide sequences. A ribozyme is an RNA molecule that catalytically cleaves other RNA molecules. Different kinds of ribozymes have been described, including group I ribozymes, hammerhead ribozymes, hairpin ribozymes, RNase P, and axhead ribozymes (see, e.g., Castanotto, et al. (1994) Adv. in Pharmacology 10 25: 289-317 for a general review of the properties of different ribozymes).

The general features of hairpin ribozymes are described, e.g., in Hampel, et al. (1990) Nucl. Acids Res. 18:299-304; European Patent Publication No. 0 360 257; U.S. Patent No. 5,254,678. Methods of preparing are well known to those of skill in the art. See, e.g., WO 94/26877; Ojwang, et al. (1993) Proc. Natl. Acad. Sci. USA 90:6340-6344; Yamada, et al. 15 (1994) Human Gene Therapy 1:39-45; Leavitt, et al. (1995) Proc. Natl. Acad. Sci. USA 92:699-703; Leavitt, et al. (1994) Human Gene Therapy 5:1151-120; and Yamada, et al. (1994) Virology 205:121-126.

Polynucleotide modulators of prostate cancer may be introduced into a cell containing the target nucleotide sequence by formation of a conjugate with a ligand binding molecule, as described in WO 91/04753. Suitable ligand binding molecules include, but are not limited to, cell surface receptors, growth factors, other cytokines, or other ligands that bind to cell surface receptors. Preferably, conjugation of the ligand binding molecule does not substantially interfere with the ability of the ligand binding molecule to bind to its corresponding molecule or receptor, or block entry of the sense or antisense oligonucleotide or its conjugated version into the cell. Alternatively, a polynucleotide modulator of prostate cancer may be introduced into a cell containing the target nucleic acid sequence, e.g., by formation of an polynucleotide-lipid complex, as described in WO 90/10448. It is understood that the use of antisense molecules or knock out and knock in models may also be used in screening assays as discussed above, in addition to methods of treatment. 20 30

Thus, in one embodiment, methods of modulating prostate disorders, e.g., cancer in cells or organisms, are provided. In one embodiment, the methods comprise administering to

a patient, e.g., to a cell within the patient, an anti-prostate cancer antibody that reduces or eliminates the biological activity of an endogenous prostate cancer protein. Alternatively, the methods comprise administering to a cell or organism a recombinant nucleic acid encoding a prostate cancer protein. This may be accomplished in many ways. In a preferred  
5 embodiment, e.g., when the prostate cancer sequence is down-regulated in prostate cancer, such state may be reversed by increasing the amount of prostate cancer gene product in the cell. This can be accomplished, e.g., by overexpressing the endogenous prostate cancer gene or administering a gene encoding the prostate cancer sequence, using known gene-therapy techniques, e.g.. In a preferred embodiment, the gene therapy techniques include the  
10 incorporation of the exogenous gene using enhanced homologous recombination (EHR), e.g., as described in PCT/US93/03868, hereby incorporated by reference in its entirety. Alternatively, e.g., when the prostate cancer sequence is up-regulated in prostate cancer, the activity of the endogenous prostate cancer gene is decreased, e.g., by the administration of a prostate cancer antisense nucleic acid.

15 In one embodiment, the prostate cancer proteins of the present invention may be used to generate polyclonal and monoclonal antibodies to prostate cancer proteins. Similarly, the prostate cancer proteins can be coupled, using standard technology, to affinity chromatography columns. These columns may then be used to purify prostate cancer antibodies useful for production, diagnostic, or therapeutic purposes. In a preferred  
20 embodiment, the antibodies are generated to epitopes unique to a prostate cancer protein; that is, the antibodies show little or no cross-reactivity to other proteins. The prostate cancer antibodies may be coupled to standard affinity chromatography columns and used to purify prostate cancer proteins. The antibodies may also be used as blocking polypeptides, as outlined above, since they will specifically bind to the prostate cancer protein.

25

#### Methods of identifying variant prostate cancer-associated sequences

Without being bound by theory, expression of various prostate cancer sequences is correlated with prostate cancer or other prostate disorders. Accordingly, disorders based on mutant or variant prostate cancer genes may be determined. In one embodiment, the  
30 invention provides methods for identifying cells containing variant prostate cancer genes, e.g., determining all or part of the sequence of at least one endogenous prostate cancer genes in a cell. This may be accomplished using many sequencing techniques. In a preferred

embodiment, the invention provides methods of identifying the prostate cancer genotype of an individual, e.g., determining all or part of the sequence of at least one prostate cancer gene of the individual. This is generally done in at least one tissue of the individual, and may include the evaluation of a number of tissues or different samples of the same tissue. The method may include comparing the sequence of the sequenced prostate cancer gene to a known prostate cancer gene, e.g., a wild-type gene.

The sequence of all or part of the prostate cancer gene can then be compared to the sequence of a known prostate cancer gene to determine if differences exist. This can be done using many known homology programs, such as Bestfit, etc. In a preferred embodiment, the presence of a difference in the sequence between the prostate cancer gene of the patient and the known prostate cancer gene correlates with a disease state or a propensity for a disease state, as outlined herein.

In a preferred embodiment, the prostate cancer genes are used as probes to determine the number of copies of the prostate cancer gene in the genome.

In another preferred embodiment, the prostate cancer genes are used as probes to determine the chromosomal localization of the prostate cancer genes. Information such as chromosomal localization finds use in providing a diagnosis or prognosis in particular when chromosomal abnormalities such as translocations, and the like are identified in the prostate cancer gene locus.

#### Administration of pharmaceutical and vaccine compositions

In one embodiment, a therapeutically effective dose of a prostate cancer protein or modulator thereof, is administered to a patient. By "therapeutically effective dose" herein is meant a dose that produces effects for which it is administered. The exact dose will depend on the purpose of the treatment, and will be ascertainable by one skilled in the art using known techniques (e.g., Ansel, et al. (1992) Pharmaceutical Dosage Forms and Drug Delivery; Lieberman (1993) Pharmaceutical Dosage Forms (vols. 1-3, Dekker, ISBN 0824770846, 082476918X, 0824712692, 0824716981; Lloyd (1999) The Art, Science and Technology of Pharmaceutical Compounding Amer. Pharma. Assn.; and Pickar (1999) Dosage Calculations Thomson). Adjustments for prostate cancer degradation, systemic versus localized delivery, and rate of new protease synthesis, as well as the age, body weight, general health, sex, diet, time of administration, drug interaction, and the severity of the

condition may be necessary, and will be ascertainable with routine experimentation by those skilled in the art. U.S. Patent Application N. 09/687,576 further discloses the use of compositions and methods of diagnosis and treatment in prostate cancer is hereby expressly incorporated by reference.

5 A "patient" for the purposes of the present invention includes both humans and other animals, particularly mammals. Thus the methods are applicable to both human therapy and veterinary applications. In the preferred embodiment the patient is a mammal, preferably a primate, and in the most preferred embodiment the patient is human. The patient typically will suffer from a prostate proliferative disorder, e.g., malignant or non-malignant, and may  
10 include cancer of other related conditions or disorders.

The administration of the prostate cancer proteins and modulators thereof of the present invention can be done in a variety of ways as discussed above, including, but not limited to, orally, subcutaneously, intravenously, intranasally, transdermally, intraperitoneally, intramuscularly, intrapulmonary, vaginally, rectally, or intraocularly. In  
15 some instances, e.g., in the treatment of wounds and inflammation, the prostate cancer proteins and modulators may be directly applied as a solution or spray, or via catheter.

The pharmaceutical compositions of the present invention comprise a prostate cancer protein in a form suitable for administration to a patient. In the preferred embodiment, the pharmaceutical compositions are in a water soluble form, such as being present as  
20 pharmaceutically acceptable salts, which is meant to include both acid and base addition salts. "Pharmaceutically acceptable acid addition salt" refers to those salts that retain the biological effectiveness of the free bases and that are not biologically or otherwise undesirable, formed with inorganic acids such as hydrochloric acid, hydrobromic acid, sulfuric acid, nitric acid, phosphoric acid and the like, and organic acids such as acetic acid, propionic acid, glycolic acid, pyruvic acid, oxalic acid, maleic acid, malonic acid, succinic  
25 acid, fumaric acid, tartaric acid, citric acid, benzoic acid, cinnamic acid, mandelic acid, methanesulfonic acid, ethanesulfonic acid, p-toluenesulfonic acid, salicylic acid, and the like. "Pharmaceutically acceptable base addition salts" include those derived from inorganic bases such as sodium, potassium, lithium, ammonium, calcium, magnesium, iron, zinc, copper,  
30 manganese, aluminum salts, and the like. Particularly preferred are the ammonium, potassium, sodium, calcium, and magnesium salts. Salts derived from pharmaceutically acceptable organic non-toxic bases include salts of primary, secondary, and tertiary amines,



substituted amines including naturally occurring substituted amines, cyclic amines and basic ion exchange resins, such as isopropylamine, trimethylamine, diethylamine, triethylamine, tripropylamine, and ethanolamine.

The pharmaceutical compositions may also include one or more of the following:  
5 carrier proteins such as serum albumin; buffers; fillers such as microcrystalline cellulose, lactose, corn and other starches; binding agents; sweeteners and other flavoring agents; coloring agents; and polyethylene glycol.

The pharmaceutical compositions can be administered in a variety of unit dosage forms depending upon the method of administration. For example, unit dosage forms  
10 suitable for oral administration include, but are not limited to, powder, tablets, pills, capsules and lozenges. It is recognized that prostate cancer protein modulators (e.g., antibodies, antisense constructs, ribozymes, small organic molecules, etc.) when administered orally, should be protected from digestion. This is typically accomplished either by complexing the molecule(s) with a composition to render it resistant to acidic and enzymatic hydrolysis, or by  
15 packaging the molecule(s) in an appropriately resistant carrier, such as a liposome or a protection barrier. Means of protecting agents from digestion are well known in the art.

The compositions for administration will commonly comprise a prostate cancer protein modulator dissolved in a pharmaceutically acceptable carrier, preferably an aqueous carrier. A variety of aqueous carriers can be used, e.g., buffered saline and the like. These  
20 solutions are typically sterile and generally free of undesirable matter. These compositions may be sterilized by conventional, well known sterilization techniques. The compositions may contain pharmaceutically acceptable auxiliary substances as required to approximate physiological conditions such as pH adjusting and buffering agents, toxicity adjusting agents and the like, e.g., sodium acetate, sodium chloride, potassium chloride, calcium chloride,  
25 sodium lactate, and the like. The concentration of active agent in these formulations can vary widely, and will be selected primarily based on fluid volumes, viscosities, body weight, and the like in accordance with the particular mode of administration selected and the patient's needs (e.g., (1980) Remington's Pharmaceutical Science (15th ed.); and Hardman, et al. (eds. 2001) Goodman & Gilman: The Pharmacological Basis of Therapeutics McGraw-Hill.

Thus, a typical pharmaceutical composition for intravenous administration would be  
30 about 0.1 to 10 mg per patient per day. Dosages from 0.1 up to about 100 mg per patient per day may be used, particularly when the drug is administered to a secluded site and not into

the blood stream, such as into a body cavity or into a lumen of an organ. Substantially higher dosages are possible in topical administration. Actual methods for preparing parenterally administrable compositions will be known or apparent to those skilled in the art, e.g., Remington's Pharmaceutical Science and Goodman and Gilman: The Pharmacological Basis of Therapeutics, supra.

The compositions containing modulators of prostate cancer proteins can be administered for therapeutic or prophylactic treatments. In therapeutic applications, compositions are administered to a patient suffering from a disease (e.g., a cancer) in an amount sufficient to cure or at least partially retard or arrest the disease and its complications. An amount adequate to accomplish this is defined as a "therapeutically effective dose." Amounts effective for this use will depend upon the severity of the disease and the general state of the patient's health. Single or multiple administrations of the compositions may be administered depending on the dosage and frequency as required and tolerated by the patient. The composition should provide a sufficient quantity of the agents of this invention to effectively treat the patient. An amount of modulator that is capable of preventing or slowing the development of cancer in a mammal is referred to as a "prophylactically effective dose." The particular dose required for a prophylactic treatment will depend upon the medical condition and history of the mammal, the particular cancer being prevented, as well as other factors such as age, weight, gender, administration route, efficiency, etc. Such prophylactic treatments may be used, e.g., in a mammal who has previously had cancer to prevent a recurrence of the cancer, or in a mammal who is suspected of having a significant likelihood of developing cancer, e.g., based partly on gene expression profiles.

It will be appreciated that the present prostate cancer protein-modulating compounds can be administered alone or in combination with additional prostate cancer modulating compounds or with other therapeutic agent, e.g., other anti-cancer agents or treatments.

In numerous embodiments, one or more nucleic acids, e.g., polynucleotides comprising nucleic acid sequences set forth in Tables 1A-4 such as antisense polynucleotides, silencing RNA, or ribozymes, will be introduced into cells, in vitro or in vivo. The present invention provides methods, reagents, vectors, and cells useful for expression of prostate cancer-associated polypeptides and nucleic acids using in vitro (cell-free), ex vivo or in vivo (cell or organism-based) recombinant expression systems.

The particular procedure used to introduce the nucleic acids into a host cell for expression of a protein or nucleic acid is application specific. Many procedures for introducing foreign nucleotide sequences into host cells may be used. These include the use of calcium phosphate transfection, spheroplasts, electroporation, liposomes, microinjection, plasma vectors, viral vectors, and many other well known methods for introducing cloned genomic DNA, cDNA, synthetic DNA, or other foreign genetic material into a host cell (see, e.g., Berger and Kimmel (1987) Guide to Molecular Cloning Techniques from Methods in Enzymology (vol. 152) Academic Press; Ausubel, et al., (eds. supplemented through 1999) Current Protocols Lippincott; and Sambrook, et al. (1989) Molecular Cloning: A Laboratory Manual (2d ed., Vol. 1-3) CSH Press.

In a preferred embodiment, prostate cancer proteins and modulators are administered as therapeutic agents, and can be formulated as outlined above. Similarly, prostate cancer genes (including both the full-length sequence, partial sequences, or regulatory sequences of the prostate cancer coding regions) can be administered in a gene therapy application. These prostate cancer genes can include antisense applications, either as gene therapy (i.e., for incorporation into the genome) or as antisense compositions, as will be appreciated by those in the art.

Prostate cancer polypeptides and polynucleotides can also be administered as vaccine compositions to stimulate HTL, CTL, and antibody responses.. Such vaccine compositions can include, e.g., lipidated peptides (see, e.g., Vitiello, et al. (1995) J. Clin. Invest. 95:341-349), peptide compositions encapsulated in poly(DL-lactide-co-glycolide) ("PLG") microspheres (see, e.g., Eldridge, et al. (1991) Molec. Immunol. 28:287-294; Alonso, et al. (1994) Vaccine 12:299-306; Jones, et al. (1995) Vaccine 13:675-681), peptide compositions contained in immune stimulating complexes (ISCOMS) (see, e.g., Takahashi, et al. (1990) Nature 344:873-875; Hu, et al. (1998) Clin Exp Immunol. 113:235-243), multiple antigen peptide systems (MAPs) (see, e.g., Tam (1988) Proc. Natl. Acad. Sci. USA 85:5409-5413; Tam (1996) J. Immunol. Methods 196:17-32), peptides formulated as multivalent peptides; peptides for use in ballistic delivery systems, typically crystallized peptides, viral delivery vectors (Perkus, et al., p. 379, in Kaufmann (ed. 1996) Concepts in vaccine development de Gruyter; Chakrabarti, et al. (1986) Nature 320:535-537; Hu, et al. (1986) Nature 320:537-540; Kieny, et al. (1986) AIDS Bio/Technology 4:790-xxx; Top, et al. (1971) J. Infect. Dis. 124:148-154; Chanda, et al. (1990) Virology 175:535-547), particles of viral or synthetic

origin (see, e.g., Kofler, et al. (1996) J. Immunol. Methods 192:25-35; Eldridge, et al. (1993) Sem. Hematol. 30:16-24; Falo, et al. (1995) Nature Med. 7:649-653), adjuvants (Warren, et al. (1986) Annu. Rev. Immunol. 4:369-388; Gupta, et al. (1993) Vaccine 11:293-306), liposomes (Reddy, et al. (1992) J. Immunol. 148:1585-1589; Rock (1996) Immunol. Today 17:131-137), or, naked or particle absorbed cDNA (Ulmer, et al. (1993) Science 259:1745-1749; Robinson, et al. (1993) Vaccine 11:957-960; Shiver, et al., p. 423, in Kaufmann (ed. 1996) Concepts in Vaccine Development de Gruyter; Cease and Berzofsky (1994) Annu. Rev. Immunol. 12:923-989; and Eldridge, et al. (1993) Sem. Hematol. 30:16-24). Toxin-targeted delivery technologies, also known as receptor mediated targeting, such as those of Avant Immunotherapeutics, Inc. (Needham, Massachusetts) may also be used.

Vaccine compositions often include adjuvants. Many adjuvants contain a substance designed to protect the antigen from rapid catabolism, such as aluminum hydroxide or mineral oil, and a stimulator of immune responses, such as lipid A, Bortadella pertussis or Mycobacterium tuberculosis derived proteins. Certain adjuvants are commercially available as, e.g., Freund's Incomplete Adjuvant and Complete Adjuvant (Difco Laboratories, Detroit, MI); Merck Adjuvant 65 (Merck and Company, Inc., Rahway, NJ); AS-2 (SmithKline Beecham, Philadelphia, PA); aluminum salts such as aluminum hydroxide gel (alum) or aluminum phosphate; salts of calcium, iron or zinc; an insoluble suspension of acylated tyrosine; acylated sugars; cationically or anionically derivatized polysaccharides; polyphosphazenes; biodegradable microspheres; monophosphoryl lipid A, and quil A. Cytokines, such as GM-CSF, interleukin-2, -7, -12, and other like growth factors, may also be used as adjuvants.

Vaccines can be administered as nucleic acid compositions wherein DNA or RNA encoding one or more of the polypeptides, or a fragment thereof, is administered to a patient. This approach is described, for instance, in Wolff, et al. (1990) Science 247:1465-1468 as well as U.S. Patent Nos. 5,580,859; 5,589,466; 5,804,566; 5,739,118; 5,736,524; 5,679,647; WO 98/04720; and in more detail below. Examples of DNA-based delivery technologies include "naked DNA", facilitated (bupivacaine, polymers, peptide-mediated) delivery, cationic lipid complexes, and particle-mediated ("gene gun") or pressure-mediated delivery (see, e.g., U.S. Patent No. 5,922,687).

For therapeutic or prophylactic immunization purposes, the peptides of the invention can be expressed by viral or bacterial vectors. Examples of expression vectors include

attenuated viral hosts, such as vaccinia or fowlpox. This approach involves the use of vaccinia virus, e.g., as a vector to express nucleotide sequences that encode prostate cancer polypeptides or polypeptide fragments. Upon introduction into a host, the recombinant vaccinia virus expresses the immunogenic peptide, and thereby elicits an immune response.

- 5 Vaccinia vectors and methods useful in immunization protocols are described in, e.g., U.S. Patent No. 4,722,848. Another vector is BCG (Bacille Calmette Guerin). BCG vectors are described in Stover, et al. (1991) *Nature* 351:456-460. A wide variety of other vectors useful for therapeutic administration or immunization, e.g., adeno and adeno-associated virus vectors, retroviral vectors, Salmonella typhi vectors, detoxified anthrax toxin vectors, and the like, will be apparent to those skilled in the art from the description herein (see, e.g., Shata, et al. (2000) *Mol. Med. Today* 6:66-71; Shedlock, et al. (2000) *J. Leuk. Biol.* 68:793-806; Hipp, et al. (2000) *In Vivo* 14:571-85).

- Methods for the use of genes as DNA vaccines are well known, and include placing a prostate cancer gene or portion of a prostate cancer gene under the control of a regulatable promoter or a tissue-specific promoter for expression in a prostate cancer patient. The prostate cancer gene used for DNA vaccines can encode full-length prostate cancer proteins, but more preferably encodes portions of the prostate cancer proteins including peptides derived from the prostate cancer protein. In one embodiment, a patient is immunized with a DNA vaccine comprising a plurality of nucleotide sequences derived from a prostate cancer gene. For example, prostate cancer-associated genes or sequence encoding subfragments of a prostate cancer protein are introduced into expression vectors and tested for their immunogenicity in the context of Class I MHC and an ability to generate cytotoxic T cell responses. This procedure may provide for production of cytotoxic T lymphocyte responses against cells which present antigen, including intracellular epitopes.

- 25 In a preferred embodiment, the DNA vaccines include a gene encoding an adjuvant molecule with the DNA vaccine. Such adjuvant molecules include cytokines that increase the immunogenic response to the prostate cancer polypeptide encoded by the DNA vaccine. Additional or alternative adjuvants are available.

- 30 In another preferred embodiment prostate cancer genes find use in generating animal models of prostate cancer. When the prostate cancer gene identified is repressed or diminished in cancer tissue, gene therapy technology, e.g., wherein antisense RNA directed to the prostate cancer gene will also diminish or repress expression of the gene. Animal

models of prostate cancer find use in screening for modulators of a prostate cancer-associated sequence or modulators of prostate cancer. Similarly, transgenic animal technology including gene knockout technology, e.g., as a result of homologous recombination with an appropriate gene targeting vector, will result in the absence or increased expression of the prostate cancer protein. When desired, tissue-specific expression or knockout of the prostate cancer protein may be necessary.

It is also possible that the prostate cancer protein is overexpressed in prostate cancer. As such, transgenic animals can be generated that overexpress the prostate cancer protein. Depending on the desired expression level, promoters of various strengths can be employed to express the transgene. Also, the number of copies of the integrated transgene can be determined and compared for a determination of the expression level of the transgene. Animals generated by such methods find use as animal models of prostate cancer and are additionally useful in screening for modulators to treat prostate cancer.

#### Kits for Use in Diagnostic and/or Prognostic Applications

For use in diagnostic, research, and therapeutic applications suggested above, kits are also provided by the invention. In the diagnostic and research applications such kits may include one of the following: assay reagents, buffers, prostate cancer-specific nucleic acids or antibodies, hybridization probes and/or primers, antisense polynucleotides, silencing RNA, ribozymes, dominant negative prostate cancer polypeptides or polynucleotides, small molecules inhibitors of prostate cancer-associated sequences, etc. A therapeutic product may include sterile saline or another pharmaceutically acceptable emulsion and suspension base.

In addition, the kits may include instructional materials containing instructions (i.e., protocols) for the practice of the methods of this invention. While the instructional materials typically comprise written or printed materials they are not limited to such. A medium capable of storing such instructions and communicating them to an end user is contemplated by this invention. Such media include, but are not limited to electronic storage media (e.g., magnetic discs, tapes, cartridges, chips), optical media (e.g., CD ROM), and the like. Such media may include addresses to internet sites that provide such instructional materials.

The present invention also provides for kits for screening for modulators of prostate cancer-associated sequences. Such kits can be prepared from readily available materials and reagents. For example, such kits can comprise one or more of the following materials: a

WO 02/098358

PCT/US02/17594

prostate cancer-associated polypeptide or polynucleotide, reaction tubes, and instructions for testing prostate cancer-associated activity. Optionally, the kit contains biologically active prostate cancer protein. A wide variety of kits and components can be prepared according to the present invention, depending upon the intended user of the kit and the particular needs of the user. Diagnosis would typically involve evaluation of a plurality of genes or products.

5 The genes will be selected based on correlations with important parameters in disease which may be identified in historical or outcome data.

## EXAMPLES

## Example 1: Gene Chip Analyses of Expression Profiles

Molecular profiles of various normal and cancerous tissues were determined and analyzed using gene chips. RNA was isolated and gene chip analysis was performed as described (Glynn, et al. (2000) *Nature* 403:672-676; Zhao, et al. (2000) *Genes Dev.* 14:981-993).

## EXAMPLE 2: Identification of androgen dependent/independent genes

To identify gene expression changes during the transition from androgen-dependent to androgen-independent prostate cancer, oligonucleotide microarrays ("K" chips or Affymetrix Eos Hu03) were interrogated with cRNAs derived from the human CWR22 prostate cancer xenograft model propagated in nude mice (Pretlow, et al. (1993) *J. Natl. Cancer Inst.* 85:394-398). The CWR22 xenograft is androgen-dependent when grown in male Nude mice. Androgen-independent sub-lines can be derived by first establishing androgen-dependent tumors in male mice. The mice are then castrated to remove the primary source of growth stimulus (androgen), resulting in tumor regression. Within 3-10 months molecular events prompt the tumors to relapse and start growing as androgen-independent tumors. See, e.g., Nagabhushan, et al. (1996) *Cancer Res.* 56:3042-3046; Amler, et al. (2000) *Cancer Res.* 60:6134-6141; and Bubendorf, et al. (1999) *J. Natl. Cancer Inst.* 91:1758-1764.

Using the CWR22 xenograft model, tumors were grown subcutaneously in male nude mice. Tumors were harvested at different times after castration. The time points post-castration included (in days): 0, 1, 3, 4, 5, 10, 30, 40, 50, 51, 52, 59, 60, 61, 70, 79, 80, 82, 120, and 125. Analyses also included established androgen-independent xenografts. Castration resulted in tumor regression. At day 120 and thereafter, the tumors relapsed and started growing in the absence of androgen.

cRNAs were generated by in vitro transcription assays (IVTs) from the different samples and were hybridized to the oligonucleotide microarrays (Affymetrix Eos Hu03). Hybridization was measured by the average fluorescence intensity (AI), which is directly proportional to the expression level of the gene.

Two types of analyses were applied to the results:

Analysis A:



The samples were divided into different time groups which included the following time points post castration (in days): 1-5, 10, 30-40, 50-82, 120-125. To identify changes in gene expression, the following calculations were made:

1. The median (or mean, in case there were only 2 samples in a group) was calculated for each group.
2. The medians (or means) for each group was compared to one-another.
3. Genes were selected that exhibited a minimum 2 fold difference in the median (or mean) between any of the groups.
4. The change in gene expression over time was analyzed for each selected gene to look for specific pattern changes.

Only genes with an interesting expression pattern during the androgen-ablation time course were selected as potential new therapeutic targets and/or diagnostic markers. Among the 70,000 gene clusters present on Hu01 and Hu02, we identified 820 gene clusters with the desired expression patterns. These expression patterns can be broadly defined into the

following categories:

1. Genes that are expressed early in the time course, then drop off in expression, and then express again with emergence of androgen-independence (hi-lo-hi pattern in Table 1A).
2. Genes that are expressed early in the time course, then drop off in expression, and do not express again with emergence of androgen-independence (hi-lo-lo pattern in Table 1A).
3. Genes that are not expressed early in the time course, but express only with emergence of androgen-independence (lo-lo-hi pattern in Table 1A).
4. Genes that are not expressed early in the time course, but then express as androgen is withdrawn and continue to express with emergence of androgen-independence (lo-hi-hi pattern in Table 1A).
5. Genes that are not expressed early in the time course, but then express as androgen is withdrawn and drop off again with emergence of androgen-independence (lo-hi-lo pattern in Table 1A).

Group 1 is characterized by cell-cycle regulating genes, such as those encoding cyclin B1, p21/WAF1, CDC18-homolog, cyclin A2, cyclin D1, and possible growth factors such as hAG2 (anterior gradient 2 homolog) among others. This indicates that interruption of growth factor and/or cell cycle pathways prevents the emergence of androgen-independent disease, making group 1 genes good targets for treating advanced prostate cancer.

Group 2 represents genes that are androgen-dependent, and do not re-express due to the lack of androgen signal in the androgen-independent phenotype. This group includes genes encoding proteins such as Fibronectin 1, which has been previously shown to be down-regulated with androgen-withdrawal (Amler, et al. (2000) Cancer Res. 60:6134-6141).

5       Group 3 represents genes that are up-regulated by signals that induce the androgen-independent phenotype. This group includes genes encoding stanniocalcin 2, c-fos proto-oncogene product, vascular endothelial growth factor, the cell surface protein transmembrane 4 superfamily member 1 and adrenomedullin among others. Adrenomedullin has recently  
10       been shown to act as an autocrine growth factor for the androgen-independent prostate cancer cell line DU145 (Rocchi, et al. (2001) Cancer Res. 61:1196-1206), indicating that its up-regulation is critical for supporting an androgen-independent phenotype. Blocking adrenomedullin function, and/or other genes in this group, prevents the growth of androgen-independent tumor cells.

15       Group 4 represents genes that are androgen-repressed and are only expressed in the absence of androgen. This group includes genes encoding the protein tyrosine phosphatase interacting protein liprin-alpha 2, the CD24 antigen, and the catalytic subunit for phosphatidylinositol 4-kinase amongst others. Patients that are treated for advanced prostate cancer by hormone-ablation may have in their bodies cells that have survived hormone-ablation and are likely to up-regulate genes that belong to Group 4. Therefore, Group 4 gene  
20       products are particularly good therapeutic targets for treating patients undergoing hormone-ablation therapy.

      Group 5 represents genes that are involved in regulating signals that induce an androgen-independent phenotype. This group includes genes encoding Rab2 (a Ras-like G protein), the Son of Sevenless homolog (a GTP/GDP exchange factor involved in activating  
25       Ras-like proteins), and the p85 regulatory subunit for phosphoinositide-3-kinase (PI3-kinase). The PI3-kinase pathway has been implicated in providing a survival signal to the prostate cancer cell line LNCaP (Lin, et al. (1999) Cancer Res. 59:2891-2897). This indicates that ras-like signals and signals dependent on PI3-kinase are involved in inducing the androgen-independent phenotype. For that reason, Group 5 gene products are particularly good  
30       therapeutic targets for treating patients undergoing hormone-ablation therapy.

Analysis B:

For the second analysis, the samples were divided into 4 time groups which included the following time points post castration (in days): 0-1, 3-5, 10-82, >120. To identify changes in gene expression, the following analysis was performed:

1. Genes were selected that exhibited a minimum of 100 AI units at the 90<sup>th</sup> percentile expression level of samples.
2. The group mean expression levels for each gene were calculated. The genes were further sub-selected to exhibit a minimum 3 fold difference between the group means.
3. An analysis of variance was then performed on selected genes. From the original 59,680 gene clusters present on the Hu03 gene chip, only about 1165 genes with a P value of < 0.01 were identified that also exhibited the above mentioned parameters.
4. A method was then employed for calculating the positive false discovery rate (pFDR), i.e., an estimate of the proportion of false-positives present in a set of findings (Storey and Tibshirani (2001) Technical Report, Department of Statistics, Stanford University, CA). This technique was developed explicitly for use with microarray data. The procedure involves randomly assigning the membership status of each sample to a group and re-performing the analysis of variance. In each simulation, the number of group members (6 for Group 1, 9 for group 2, 15 for group 3, and 4 for group 4) remained constant, but these designations were shuffled and assigned to each sample at random. The permutation was performed 1000 times, and for each simulation, the number of findings at  $P < 0.01$  was noted. The number of false positives under null conditions, was then divided by the number of actual findings ( $n=1165$  genes) to obtain an estimate of the proportion of false positive findings. After the application of a correction factor, the final estimate for the pFDR was about 1%. Thus, one can expect that approximately 12 of the 1165 findings are false positives.
5. The approximately 1165 genes were clustered by expression pattern to identify specific pattern changes. Only genes with an interesting expression pattern during the androgen-ablation time course were selected as potential new therapeutic targets and/or diagnostic markers. These expression patterns can be broadly defined into the following categories:
  1. Genes that are expressed early in the time course of androgen withdrawal, then drop off in expression, and then express again with emergence of androgen-independence (hi-lo-lo-hi pattern in Table 2A).

2. Genes that are expressed early in the time course, then drop off in expression immediately after androgen-withdrawal, and do not express again with emergence of androgen-independence (hi-lo-lo-lo pattern in Table 2A).
3. Genes that are expressed early in the time course, then drop off in expression after several  
5 days of androgen withdrawal, and do not express again with emergence of androgen-independence (hi-hi-lo-lo pattern in Table 2A).
4. Genes that are not expressed early in the time course, but express only with emergence of androgen-independence (lo-lo-lo-hi pattern in Table 2A).
5. Genes that are not expressed early in the time course, but then express as androgen is  
10 withdrawn and continue to express with emergence of androgen-independence (lo-lo-hi-hi pattern in Table 2A).
6. Genes that are not expressed early in the time course, but then express as androgen is withdrawn and drop off again with emergence of androgen-independence (lo-lo-hi-lo pattern in Table 2A).

15       Group 1 is characterized by cell-cycle regulating genes and cell growth promoting genes, such as those encoding cyclin B1 and CDC45 among others, growth factors/hormones such as hAG2 (anterior gradient 2 homolog), adrenomedullin, and stanniocalcin 2 among others, and growth factor receptors, such as the bone morphogenic protein receptor type 1B (BMP-R1B) and the endothelial differentiation lysophosphatidic acid G-protein-coupled  
20       receptor 7 among others. Adrenomedullin has recently been shown to act as an autocrine growth factor for the androgen-independent prostate cancer cell line DU145 (Rocchi, et al. (2001) Cancer Res. 61:1196-1206), indicating that its up-regulation is critical for supporting an androgen-independent phenotype. This indicates that interruption of growth factor and/or cell cycle pathways prevents the emergence of androgen-independent disease, making group  
25       1 genes good targets for treating both localized and advanced prostate cancer and related conditions.

      Group 2 represents genes that are androgen-dependent, and do not re-express due to the lack of androgen signal in the androgen-independent phenotype. This group includes  
genes encoding proteins such as the endothelial protein C receptor (EPCR) and the potassium  
30       intermediate/small conductance calcium-activated channel (subfamily N, member 2). These genes represent targets for treating androgen-dependent prostate cancer and related conditions.

Group 3 also represents genes that are androgen-dependent, and do not re-express due to the lack of androgen signal in the androgen-independent phenotype. This group includes genes encoding proteins such as Fibronectin 1, which has been previously shown to be down-regulated with androgen-withdrawal (Amler, et al. (2000) Cancer Res. 60:6134-6141), and  
5 genes encoding signaling proteins such as Rho GTPase activating protein 1. These genes represent targets for treating androgen-dependent prostate cancer and related conditions.

Group 4 represents genes that are up-regulated by signals that induce and maintain the androgen-independent phenotype. This group includes genes encoding potential growth promoting proteins such as chemokine-like factor (Unigene ID Hs.15159), colon cancer-associated protein Mic1, and the mitogen-activated protein kinase-activated protein kinase.  
10 Blocking function of these proteins, and/or other genes in this group, prevents the growth of androgen-independent tumor cells and related conditions.

Group 5 represents genes that are androgen-repressed and are only expressed in the absence of androgen or that are induced by the absence of androgen. This group includes  
15 genes encoding transcriptional regulators such as the androgen receptor, the DNA activated protein kinase (catalytic subunit), and nuclear factor related to kappa B binding protein (NFRKB), among others. Patients that are treated for advanced prostate cancer by hormone-ablation may have in their bodies cells that have survived hormone-ablation and are likely to up-regulate genes that belong to Group 5. Therefore, Group 5 gene products are particularly  
20 good therapeutic targets for treating patients undergoing hormone-ablation therapy.

Group 6 represents genes that are involved in regulating signals that are induced during androgen withdrawal and that induce an androgen-independent phenotype. This group includes genes encoding signaling molecules such as phosphoinositide-3-kinase (class 2, alpha polypeptide), signal transducer and activator of transcription 2 (STAT2), phospholipase  
25 A2 (group IIA) and the protein tyrosine phosphatase interacting protein liprin-alpha 2, cell surface receptors such as gamma-aminobutyric acid (GABA) A receptor epsilon subunit, G protein-coupled receptor 48, and immune function proteins such as the major histocompatibility complex class II DR alpha. The PI3-kinase pathway has been implicated in providing a survival signal to the prostate cancer cell line LNCaP (Lin, et al. (1999) Cancer  
30 Res. 59:2891-2897). This indicates that ras-like signals and signals dependent on PI3-kinase are involved in inducing the androgen-independent phenotype. For that reason, Group 6 gene

WO 02/098358

PCT/US02/17594

products are particularly good therapeutic targets for treating patients undergoing hormone-ablation therapy.

TABLE 1A provides Accession numbers for genes, including expressed sequence tags, incorporated in their entirety here and throughout the application where Accession numbers are provided). Genes with an interesting expression pattern during the androgen-ablation time course were selected as potential new therapeutic targets and/or diagnostic markers. 820 gene clusters were identified with desired expression patterns. These expression patterns can be broadly defined into the following categories:

1. Genes that are expressed early in the time course, then drop off in expression, and then express again with emergence of androgen-independence (hi-to-lo pattern).
2. Genes that are expressed early in the time course, then drop off in expression, and do not express again with emergence of androgen-independence (hi-to-lo pattern).
3. Genes that are not expressed early in the time course, but express only with emergence of androgen-independence (lo-to-hi pattern).
4. Genes that are not expressed early in the time course, but then express as androgen is withdrawn and continue to express with emergence of androgen-independence (lo-to-hi pattern).
5. Genes that are not expressed early in the time course, but then express as androgen is withdrawn and drop off again with emergence of androgen-independence (lo-to-lo pattern).

Table 1B lists accession numbers for primers lacking a unique/D in table 1A. For each probe/ primer is listed a gene cluster number from which oligonucleotides were designed. Gene clusters were compiled using sequences derived from Genbank ESTs and mRNAs. These sequences were clustered based on sequence similarity using Clustering and Alignment Tools (Doublet/visi, Oakland California). Genbank accession numbers for sequences comprising each cluster are listed in the "Accession" column.

Table 1C lists genomic positioning for primers lacking unique/Ds and accession numbers in tables 1A. For each predicted exon is listed genomic sequence source used for prediction. Nucleotide locations of each predicted exon are also listed.

TABLE 1A

20	Play	Exon	Unigene/D	Unigene Title	pattern
	102772	U83115	Hs.151002	absent in melanoma 1	hi-to-hi
	128610	M48373	Hs.10247	activated leucocyte cell adhesion molecule	hi-to-hi
	102276	M48373	Hs.10247	activated leucocyte cell adhesion molecule	hi-to-hi
25	100554	A03758			hi-to-hi
	100555	A03758			hi-to-hi
	135400	XT8182	Hs.59915	androgen receptor (dihydrotestosterone r	hi-to-hi
	331363	AW582256	"Hs.91011	anterior gradient 2 (Xenopus laevis) hom	hi-to-hi
	115764	AW582256	"Hs.91011	anterior gradient 2 (Xenopus laevis) hom	hi-to-hi
30	120483	BE251623	Hs.1578	baculoviral IAP repeat-containing 5 (bar	hi-to-hi
	101605	AA502860	Hs.75892	asparagine synthetase	hi-to-hi
	127236	AW561857	Hs.98568	binding uninhibited by benzimidazole 1	hi-to-hi
	128472	BE241980	"Hs.10029	calnexin C	hi-to-hi
	102712	U77949	Hs.65662	CD66 (cell division cycle 6, S. cerevisiae)	hi-to-hi
	314843	YU0272	Hs.184572	cell division cycle 2, G1 to S and G2 to	hi-to-hi
	102122	NM_001809	Hs.1594	centromere protein A (CTM)	hi-to-hi
	325213			CH17_hs_g9597224	hi-to-hi
	327110			CH21_hs_g9117842	hi-to-hi
40	339186			CH22_D459H16.GENSCAN.72-13	hi-to-hi
	337755			CH22_EMAC000097.GENSCAN.109-2	hi-to-hi
	337674			CH22_EMAC000097.GENSCAN.87-4	hi-to-hi
	337675			CH22_EMAC000097.GENSCAN.87-6	hi-to-hi
	332616			CH22_FGENES.173_1	hi-to-hi
	333517			CH22_FGENES.173_2	hi-to-hi
45	333796			CH22_FGENES.275_1	hi-to-hi
	333796			CH22_FGENES.275_3	hi-to-hi
	333808			CH22_FGENES.275_2	hi-to-hi
	333809			CH22_FGENES.280_2	hi-to-hi
	333792			CH22_FGENES.3_2	hi-to-hi
50	334101			CH22_FGENES.277_59	hi-to-hi
	334502			CH22_FGENES.297_18	hi-to-hi
	334616			CH22_FGENES.411_15	hi-to-hi
	334859			CH22_FGENES.452_13	hi-to-hi
	334900			CH22_FGENES.452_14	hi-to-hi
55	334902			CH22_FGENES.452_16	hi-to-hi
	334905			CH22_FGENES.452_20	hi-to-hi
	334906			CH22_FGENES.452_21	hi-to-hi
	334951			CH22_FGENES.465_20	hi-to-hi
	335044			CH22_FGENES.480_1	hi-to-hi
60	335753			CH22_FGENES.604_2	hi-to-hi
	335755			CH22_FGENES.604_4	hi-to-hi
	333135			CH22_FGENES.63_11	hi-to-hi
	333137			CH22_FGENES.63_13	hi-to-hi
	333138			CH22_FGENES.63_16	hi-to-hi
65	333139			CH22_FGENES.63_16	hi-to-hi
	336721			CH22_FGENES.63-17	hi-to-hi
	105012	AF098158	Hs.9329	chromosome 20 open reading frame 1	hi-to-hi
	134470	X54942	Hs.83758	CDG8 protein kinase 2	hi-to-hi
70	134750	L29073	Hs.1139	cold shock domain protein A	hi-to-hi
	125919	AA044842	"Hs.251671	C1P synthase	hi-to-hi
	102993	BE262939	Hs.85137	cyclin A2	hi-to-hi
	131185	BE280074	Hs.23960	cyclin B1	hi-to-hi
	106350	AK001034	"Hs.194698	cyclin D2	hi-to-hi
75	103380	AI077231	"Hs.82332	cyclin D1 (PRAD1; parathyroid adenomas	hi-to-hi
	101216	AA284195	Hs.84113	cyclin-dependent kinase inhibitor 3 (CDK	hi-to-hi
	100559	AW247430	Hs.84152	cystathionine-beta-synthase	hi-to-hi
	130555	AI831062	Hs.17409	cysteine-rich protein 1 (p53nif)	hi-to-hi
	101473	M22376	Hs.83834	cyclothione b-5	hi-to-hi
80	101469	BE533926	"Hs.191003	cyclothione c oxidase subunit Va	hi-to-hi
	103546	Z14244	"Hs.75752	cyclothione c oxidase subunit Vb	hi-to-hi
	100829	AA471098	Hs.278544	acyl-CoA:acyl-CoA transferase 2 (a	hi-to-hi
	102469	AF059293	Hs.180015	D-dependent transaminase	hi-to-hi







WO 02/098358

PCT/US02/17594

	126646	AA316161	Hs.61635	six transmembrane epithelial antigen of	Hs-hi
	103059	XG7349	Hs.184510	stratifin	Hs-hi
	102632	U66618	Hs.250581	SWI/SNF related, matrix associated, acti	Hs-hi
5	103269	AF230662	*Hs.289106	syneovial sarcoma, X breakpoint 2	Hs-hi
	126820	AA622037	Hs.166468	programmed cell death 5	Hs-hi
	101114	XU3328	Hs.83642	thymidylate synthase	Hs-hi
	102846	BE264974	Hs.6566	thyroid hormone receptor interactor 13	Hs-hi
	131877	JQ4088	*Hs.156346	thymosin (DNA) $\alpha$ (170kD)	Hs-hi
10	100866	U14134	Hs.75113	general transcription factor IIA	Hs-hi
	133953	AK434699	Hs.77256	transferrin receptor (g60, CD71)	Hs-hi
	130135	AA511426	*Hs.21635	tissue, gamma 1	Hs-hi
	130267	AA478005	Hs.154036	tumor suppressing substrate protein candid	Hs-hi
	126180	L32977	Hs.3712	ubiquitin-cytochrome c reductase, Rhesus	Hs-hi
15	101536	NM_009002	Hs.77517	ubiquitin carboxy-terminal esterase L3	Hs-hi
	102867	NM_007019	Hs.83002	ubiquitin carboxy-terminal esterase	Hs-hi
	103566	Z19002	Hs.37096	zinc finger protein 145 (K-rap80-like, e	Hs-hi
	300022				Hs-hi
	133016	AJ002744	Hs.246315	UDP-N-acetyl-alpha-D-glucosaminase poly	Hs-hi
20	126642	NM_001380	Hs.11836	7-dehydrocholesterol reductase	Hs-hi
	154568	AF207664	Hs.8230	a disintegrin-like and metalloprotease {	Hs-hi
	300023				Hs-hi
	125183	AV660804	Hs.301417	AHRNAK nucleoprotein (desmoyokin)	Hs-hi
	101766	MB0889	*Hs.301417	AHRNAK nucleoprotein (desmoyokin)	Hs-hi
	133516	BE265133	*Hs.217493	annexin A2	Hs-hi
25	102148	AW162057	Hs.78628	ATPase, Na+/K+-transporting, beta 1 poly	Hs-hi
	105538	AT50879	Hs.74034	Homo sapiens clone 24651 mRNA sequence	Hs-hi
	102854	AB78826	Hs.323469	carboxyl 1, caveolin protein, Z2D0	Hs-hi
	323665			CHX-1, hG656836	Hs-hi
30	334282			CHZ2_FGENES.389_12	Hs-hi
	334891			CHZ2_FGENES.462_5	Hs-hi
	335149			CHZ2_FGENES.469_5	Hs-hi
	335682			CHZ2_FGENES.516_2	Hs-hi
	335766			CHZ2_FGENES.604_5	Hs-hi
35	303951	AW475081	Hs.172928	collagen, type I, alpha 1	Hs-hi
	134421	AU077196	Hs.82885	collagen, type V, alpha 2	Hs-hi
	131101	BE363561	Hs.22581	DMP-250811523 protein	Hs-hi
	121153	AU077333	*Hs.160483	erythrocyte membrane protein band 7.2 (s	Hs-hi
	103328	AU077333	*Hs.160483	erythrocyte membrane protein band 7.2 (s	Hs-hi
40	322035	AL137517	*Hs.306201	hypothetical protein DKFZp564O1278	Hs-hi
	301872	HB4730	Hs.326391	ESTs, Highly similar to KIAA1437 protein	Hs-hi
	302820	AE037858	Hs.173484	hypothetical protein FLJ10337	Hs-hi
	300409	T58155		glycylserine 3.1 Stratagene lung (93721) H	Hs-hi
	304735	AA578453		glucan7sh11.s1 NCL_OGAP_Cos Homo sapiens	Hs-hi
45	306956	A1138628	Hs.308059	EST, Weakly similar to zinc finger prot	Hs-hi
	126789	AW366576	Hs.133651	carboxyl 2	Hs-hi
	130267	AB037698	Hs.173484	hypothetical protein FLJ10337	Hs-hi
	114765	AE037858	Hs.173484	hypothetical protein FLJ10337	Hs-hi
	104204	AK001691	Hs.57655	hypothetical protein FLJ10629	Hs-hi
	105200	AA328102	Hs.24641	cytoskeleton associated protein 2	Hs-hi
50	105453	AL047596	Hs.10293	RNA binding motif protein 18	Hs-hi
	107677	A1188161	Hs.144627	ESTs	Hs-hi
	108880	AA766605	*Hs.47095	hypothetical protein FLJ1212	Hs-hi
	111157	AL105728	Hs.18948	ESTs, Highly similar to A15026 probable	Hs-hi
55	116202	BE168035	Hs.87089	ESTs	Hs-hi
	120689	AW134519	Hs.36125	ESTs	Hs-hi
	121847	AA648828	Hs.2759	cardiopoietic protein 1	Hs-hi
	121822	AA637471	Hs.107601	ESTs	Hs-hi
	128515	BE365805	Hs.10286	type 1 transmembrane protein Fc14	Hs-hi
60	130486	W16744	Hs.180069	Homo sapiens cDNA FLJ20655 fs, clone KA	Hs-hi
	131076	AA748230	Hs.22666	ESTs	Hs-hi
	131084	NM_017413	Hs.303084	apelin, peptide ligand for APJ receptor	Hs-hi
	134109	AA540031	Hs.7913	ESTs	Hs-hi
	300289	AW78533	Hs.186260	ESTs	Hs-hi
	302767	HS4900	Hs.17882	ESTs	Hs-hi
65	312391	RA3707	Hs.133159	ESTs, Weakly similar to PHUSD salivary	Hs-hi
	312889	AW404061	Hs.203965	ESTs	Hs-hi
	316716	AB284218	Hs.130749	ESTs	Hs-hi
	315843	AA678430	Hs.191897	ESTs	Hs-hi
	322447	A1735759	Hs.52620	Integrin, beta 8	Hs-hi
70	322826	AB07893	Hs.201771	ESTs	Hs-hi
	324667	AB24707	*Hs.55151	Homo sapiens cDNA: FLJ21682 fs, clone C	Hs-hi
	331136	AA287450	Hs.53842	Homo sapiens cDNA: FLJ22654 fs, clone	Hs-hi
	331353	AA953008	Hs.88143	ESTs	Hs-hi
	133063	AB54133	Hs.30212	thyroid receptor interacting protein 15	Hs-hi
75	311034	BE667130	Hs.211389	ESTs, Moderately similar to PT0375 natri	Hs-hi
	108647	BE546947	Hs.44276	homo box C10	Hs-hi
	124955	AA376768	*Hs.324841	hypothetical protein FLJ22622	Hs-hi
	113823	AW953484	Hs.3849	hypothetical protein FLJ22041 similar to	Hs-hi
80	313657	AK131798	Hs.164192	ESTs, Weakly similar to Y161 HUMAN HYPOT	Hs-hi
	302949	AB91344	Hs.127819	ESTs, Weakly similar to T17336 hypotheti	Hs-hi
	128453	X02761	*Hs.287820	fibronectin 1	Hs-hi
	306232	AA670052	Hs.169476	glyoxaldehyde-3-phosphate dehydrogenase	Hs-hi

WO 02/098358

PCT/US02/17594

	117642	U55184	Hs.154145	hypothetical protein FLJ11585	hi-to-hi
	116881	NM_006766	Hs.184942	G protein-coupled receptor 64	hi-to-hi
	133666	U56725	Hs.75452	heat shock 70kD protein 2	hi-to-hi
5	103282	X78565	Hs.289114	hexabrachion (hesenin C, cyclodextrin)	hi-to-hi
	107073	S69027		gb.HCX CB-class 1 homeodomain (fragment)	hi-to-hi
	102289	U32114			hi-to-hi
	319109	Z45682	Hs.30797	Homo sapiens clone Z520 mRNA sequence	hi-to-hi
	116357	AF052107	Hs.30797	Homo sapiens clone Z520 mRNA sequence	hi-to-hi
10	107457	W65150	Hs.37304	homo box A5	hi-to-hi
	105508	AA173942	Hs.335416	Homo sapiens mRNA; cDNA DKF7Z654H1016 (f)	hi-to-hi
	302290	AA173949	Hs.175563	Homo sapiens mRNA; cDNA DKF7Z654H10763 (f)	hi-to-hi
	102838	R34687	Hs.80658	uncoupling protein 2 (mitochondrial, pro)	hi-to-hi
	106235	Z29594	Hs.13421	KIAA055 protein	hi-to-hi
15	133567	NM_002206	Hs.74369	Integrin, alpha 7	hi-to-hi
	125573	AI351642	Hs.182241	Interferon induced transmembrane protein	hi-to-hi
	103059	X57351	Hs.174135	Interferon induced transmembrane protein	hi-to-hi
	330415	D83777	Hs.75137	KIAA0153 gene product	hi-to-hi
	363054	BE762848	Hs.289380	colon cancer-associated protein Mct1	hi-to-hi
20	133879	X75346	Hs.75074	mitogen-activated protein kinase-activated	hi-to-hi
	106528	BE366801	Hs.21858	nucleotide repeat containing 3	hi-to-hi
	107480	AF001691	Hs.74304	peripalin	hi-to-hi
	133050	X73424	Hs.63788	prolyl 4-hydroxylase, beta p	hi-to-hi
	133981	AI184531	Hs.208538	prostate differentiation factor	hi-to-hi
	106390	AJ297438	Hs.20168	prostate stem cell antigen	hi-to-hi
25	302124	AA676403	Hs.145078	regulator of differentiation (n. S. pom	hi-to-hi
	123823	X00949	Hs.106314	relaxin 1 (H)	hi-to-hi
	134444	BE_184455	Hs.251754	secretory leukocyte protease inhibitor 1	hi-to-hi
	103240	U81981	Hs.2794	sodium channel, nonvoltage-gated 1 alpha	hi-to-hi
30	116781	AA368037	Hs.30911	solute carrier family 16 (monocarboxylic	hi-to-hi
	321412	AI674383	Hs.22891	solute carrier family 7 (cationic amino	hi-to-hi
	125487	AA233909	Hs.184601	solute carrier family 7 (cationic amino	hi-to-hi
	107759	ME0244	Hs.184601	solute carrier family 7 (cationic amino	hi-to-hi
	112941	AW163034	Hs.6467	synaptophysin 3	hi-to-hi
35	134351	BE272506	Hs.82109	syndecan 1	hi-to-hi
	125924	BE272506	Hs.82109	syndecan 1	hi-to-hi
	133982	AJ033927	Hs.21858	nucleotide repeat containing 3	hi-to-hi
	133473	AW301933	Hs.73380	Imporin T1, skeletal, slow	hi-to-hi
	101042	T46839	Hs.10319	UDP glycosyltransferase 2 family, polype	hi-to-hi
	123955	X77777	Hs.198726	vasoactive intestinal peptide receptor 1	hi-to-hi
40	102922	BE5420	Hs.185191	villin 2 (adult)	hi-to-hi
	106858	BE185536	Hs.300815	Homo sapiens mRNA; cDNA DKF7Z654H1172 (f)	hi-to-hi
	132818	AL050025	Hs.279918	hypothetical protein FLJ20151	hi-to-hi
	100187	D17753	Hs.78183	aldo-keto reductase family 1, member C3	hi-to-hi
45	116334	AL038450	Hs.48948	ATP2C1 calcium transport ATPase, same as	hi-to-hi
	134464	NM_013320	Hs.288124	CD24 antigen (small cell lung carcinoma	hi-to-hi
	302067	BE542706	Hs.222399	CEP1 protein	hi-to-hi
	105500	AW602166	Hs.222399	CEP1 protein	hi-to-hi
	100732	AA557880	Hs.78152	decorin	hi-to-hi
50	123955	AA538992	Hs.171656	dual specificity phosphatase 1	hi-to-hi
	117789	N48294	Hs.46850	EST	hi-to-hi
	330786	BE379594	Hs.49136	ESTs, Moderately similar to ALU7_HUMAN A	hi-to-hi
	319808	T28950	Hs.17283	hypothetical protein FLJ10850	hi-to-hi
	305052	BE174240	Hs.221132	gb.GV1-HT0573-280200-062.06 HT0573 Homo	hi-to-hi
55	116780	H22688	Hs.30008	ESTs	hi-to-hi
	104189	AB040527	Hs.301804	KIAA1454 protein	hi-to-hi
	105888	L43821	Hs.80281	enhancer of filamentation 1 (cas-like)	hi-to-hi
	105731	AA343864	Hs.23131	nuclear receptor coactivator 2	hi-to-hi
	105772	HE7111	Hs.221132	ESTs	hi-to-hi
	106794	H24530	Hs.272394	hypothetical protein FLJ20069	hi-to-hi
60	113098	N77737	Hs.8349	Apobec-1 complementation factor; APOBEC-	hi-to-hi
	114530	AA601038	Hs.191787	chromosome 5 open reading frame 4	hi-to-hi
	116188	AA468183	Hs.184598	Homo sapiens cDNA, FLJ23241 f1, clone C	hi-to-hi
	117330	AI904095	Hs.43423	ESTs	hi-to-hi
65	117701	BE053921	Hs.295971	ESTs	hi-to-hi
	123911	AI189754	Hs.144330	ESTs	hi-to-hi
	124083	AW195237	Hs.7734	hypothetical protein FLJ22174	hi-to-hi
	124690	AW883529	Hs.172830	ESTs	hi-to-hi
	130796	AA088809	Hs.13025	hypothetical protein FLJ22794	hi-to-hi
70	131624	AB040527	Hs.301804	KIAA1454 protein	hi-to-hi
	132116	AW892474	Hs.40289	ESTs	hi-to-hi
	132442	AW970859	Hs.313503	ESTs	hi-to-hi
	310219	AI221087	Hs.147761	ESTs	hi-to-hi
	310398	AI439136	Hs.142546	ESTs	hi-to-hi
75	310884	AWC14684	Hs.232189	ESTs	hi-to-hi
	311587	AI328254	Hs.271019	ESTs, Weakly similar to SMNH_HUMAN SURV	hi-to-hi
	312240	R36475	Hs.24321	Homo sapiens cDNA FLJ17028 f1, clone HE	hi-to-hi
	312853	AA677934	Hs.117864	ESTs	hi-to-hi
	314219	AA292331	Hs.48378	Homo sapiens clone HB-2 mRNA sequence	hi-to-hi
80	315052	AA876910	Hs.134427	ESTs	hi-to-hi
	331919	AA448869	Hs.118316	ESTs	hi-to-hi
	133240	AF001489	Hs.242894	ADP-ribosylation factor-1-like 1	hi-to-hi

WO 02/098358

PCT/US02/17594

	134008	Z45957	Hs.7937	G-protein-coupled receptor induced prote	to-hi-hi
	124847	W07701	Hs.304177	Homo sapiens clone FL85503 PRG2286 mRNA,	to-hi-hi
	129087	A348027	Hs.106557	Homo sapiens clone PP1057 unknown mRNA	to-hi-hi
5	131762	AA744902	Hs.107787	hypothetical protein PRG1489	to-hi-hi
	123000	AA744902	Hs.107787	hypothetical protein PRG1489	to-hi-hi
	106713	A122943	Hs.194319	ESTs, Weakly similar to KIAA1006 protein	to-hi-hi
	118475	N68645		glic2a46c11.1.1 Scores fetal liver spleen	to-hi-hi
	118381	N84513	Hs.43894	ESTs, Weakly similar to AF151800 1 CGI-4	to-hi-hi
10	105057	AA194233		glic2a46c11.1.1 Straglene colon (S37204)	to-hi-hi
	131507	AB26209	Hs.27769	ESTs, Weakly similar to MCAT_HUMAN MITOC	to-hi-hi
	124970	BE272862	Hs.106534	hypothetical protein FLJ22625	to-hi-hi
	130094	NM_001471	Hs.167017	gamma-aminobutyric acid (GABA) B recepto	to-hi-hi
	302357	X03178	Hs.198246	group-specific component (vitamin D bind	to-hi-hi
15	113231	AA278583	Hs.180737	Homo sapiens clone 23864 and 23005 mRNA	to-hi-hi
	111523	BE383234	Hs.25925	Homo sapiens clone 23860 mRNA sequence	to-hi-hi
	128530	AB52595	Hs.183475	Homo sapiens clone 25051 mRNA sequence	to-hi-hi
	129997	A335048	Hs.107637	hypothetical protein FLJ12806	to-hi-hi
	315306	AB037745	Hs.134858	KIAA1324 protein	to-hi-hi
20	133844	AW08579	Hs.7790	Homo sapiens mRNA; cDNA DKFZp564A072 (fr	to-hi-hi
	115054	BE383688	Hs.42464	hypothetical protein FLJ10618	to-hi-hi
	128593	AA373514	Hs.5937	Homo sapiens mRNA; cDNA DKFZp586P1522 (f	to-hi-hi
	109523	AW07385	Hs.235901	KIAA0438 protein	to-hi-hi
	130577	MS241	Hs.142	Insulin-like growth factor binding prote	to-hi-hi
25	101859	AF185747	Hs.181350	kallikrein 3, prostatic	to-hi-hi
	130336	AA536210	Hs.171995	kallikrein 3, (prostate specific antigen	to-hi-hi
	129190	AW945088	Hs.171995	kallikrein 3, (prostate specific antigen	to-hi-hi
	134921	AL137491	Hs.125511	Homo sapiens mRNA; cDNA DKFZp434P1530 (f	to-hi-hi
	302355	AJ224172	Hs.204068	lysozin B (steroglobin family member)	to-hi-hi
30	117921	AJ021459	Hs.306490	Homo sapiens mRNA; cDNA DKFZp761E212 (f	to-hi-hi
	101701	NM_002436	Hs.1981	membrane protein, palmitoylated 1 (55kD)	to-hi-hi
	130356	AF127577	Hs.155107	nuclear receptor interacting protein 1	to-hi-hi
	101763	AB001914	Hs.170414	paired basic amino acid cleaving system	to-hi-hi
	130342	AB18102	Hs.154948	phosphatidylglycerol 4-kinase, catalytic	to-hi-hi
35	130780	AW379130	Hs.15953	phosphodiesterase 9A	to-hi-hi
	101461	N9559	Hs.79422	phospholipase A2, group 1A (phalate,	to-hi-hi
	134032	NM_005025	Hs.73919	serine (or cysteine) proteinase inhibitor	to-hi-hi
	303792	AF034799	Hs.30891	protein tyrosine phosphatase, receptor 1	to-hi-hi
	110932	AJ021459	Hs.306490	Homo sapiens mRNA; cDNA DKFZp761E212 (f	to-hi-hi
40	135182	U83993	Hs.321708	purinergic receptor P2X, ligand-gated io	to-hi-hi
	133985	U97275	Hs.177201	quiescin QSOX2	to-hi-hi
	134142	BE244253	Hs.75292	retinol oxidase-like 2 (p130)	to-hi-hi
	100977	X80921	Hs.302177	H.sapiens mRNA for ribosomal protein L19	to-hi-hi
	133534	AU077115	Hs.201675	RNA binding motif protein 5	to-hi-hi
45	133011	NM_008379	Hs.171921	serine domain, immunoglobulin domain (Ig,	to-hi-hi
	132180	W26406	Hs.258293	seven h. albicollis (Drosophila) homing 1	to-hi-hi
	103110	X82922	Hs.2554	sialyltransferase 1 (beta-galactoside al	to-hi-hi
	130173	U38847	Hs.151519	TAR (HIV) RNA-binding protein 1	to-hi-hi
	127436	X95085	Hs.289191	Homo sapiens cDNA FLJ136131a, clone PL	to-hi-hi
50	110820	N84039	Hs.4062	lecithin, galactoside-binding, soluble, 8	to-hi-hi
	114680	AJ071353		gluc2a46c10.1.1 Straglene fibroblast (S37	to-hi-hi
	330541	NM_002038	Hs.289827	interferon, alpha-inducible protein (do	to-hi-hi
	101485	AA535324	Hs.1852	acid phosphatase, prostatic	to-hi-hi
	332386	NM_004061	Hs.102	aminosulfolactonase (glycine cleavage	to-hi-hi
55	100559	AA535210	Hs.171995	kallikrein 3, (prostate specific antigen	to-hi-hi
	134738	AU079801	Hs.89436	cadherin 17, LI cadherin (fiver-intestin	to-hi-hi
	103119	X83629	Hs.2577	cadherin 3, type 1, P-cadherin (placenta	to-hi-hi
	302952	AW176909	Hs.42548	cadherin-binding protein calcisarin-1	to-hi-hi
	105402	AB014880	Hs.3706	carbohydrate (phosphatidylcholine) sul	to-hi-hi
	102978	AU077174	Hs.289191	caldesmon H	to-hi-hi
60	101783	W01076	Hs.115853	CD59 antigen p18-20 (antigen identified	to-hi-hi
	129890	AB888572	Hs.282904	Homo sapiens cDNA, FLJ22704 fls, clone H	to-hi-hi
	328164			CH16, Jn pJ558863	to-hi-hi
	328648			CH.07, Jn pJ600473	to-hi-hi
	330032			CH.16, p2 pJ6882505	to-hi-hi
65	330033			CH.16, p2 pJ6882506	to-hi-hi
	328916			CH.20, Jn pJ555658	to-hi-hi
	337603			CH22, C20H12.GENSCAN.16-2	to-hi-hi
	338581			CH22, EMAC005900.GENSCAN.421-5	to-hi-hi
70	338582			CH22, EMAC005900.GENSCAN.421-6	to-hi-hi
	333743			CH22, FGENES.264_1	to-hi-hi
	333846			CH22, FGENES.290_3	to-hi-hi
	333849			CH22, FGENES.290_8	to-hi-hi
	334221			CH22, FGENES.300_1	to-hi-hi
	334222			CH22, FGENES.300_3	to-hi-hi
75	334578			CH22, FGENES.406_1	to-hi-hi
	338682			CH22, FGENES.41-1	to-hi-hi
	338694			CH22, FGENES.46-1	to-hi-hi
	338289			CH22, FGENES.527_2	to-hi-hi
	335290			CH22, FGENES.527_3	to-hi-hi
	335293			CH22, FGENES.527_6	to-hi-hi
	337192			CH22, FGENES.570-2	to-hi-hi
80	335909			CH22, FGENES.617_6 (same as BFH)	to-hi-hi



WO 02/098358

PCT/US02/17594

	132932	AW118826	Hs.6093	Homo sapiens cDNA: FLJ22783 fls, clone K	to-hi-to
	134690	BE326276	*Hs.8801	ESTs	to-hi-to
	300967	AA565209	Hs.269439	ESTs	to-hi-to
5	301182	AW291411	Hs.152531	ESTs, Weakly similar to 500754 zinc finger	to-hi-to
	302595	A1863372	Hs.152547	Homo sapiens mRNA; cDNA DKFZp454A171 (fr	to-hi-to
	303132	A182819	Hs.4655	chromosome 21 open reading frame 50	to-hi-to
	303506	AA340605	Hs.105887	ESTs, Weakly similar to Homolog of rat Z	to-hi-to
	303654	BE246743	Hs.288529	hypothetical protein FLJ22535	to-hi-to
10	310026	AA276233	Hs.100891	ESTs	to-hi-to
	310056	A1233072	Hs.143593	ESTs	to-hi-to
	310363	A1261700	Hs.145544	ESTs	to-hi-to
	310371	A1262584	Hs.145575	ESTs	to-hi-to
	310430	A1270843	Hs.200257	ESTs	to-hi-to
15	310438	AW221162	Hs.200197	ESTs	to-hi-to
	310455	A1277603	Hs.145590	ESTs	to-hi-to
	310787	AW262580	Hs.147674	KIAA1621 protein	to-hi-to
	311067	A0587332	Hs.209115	ESTs	to-hi-to
	311422	F00677	Hs.101316	ESTs	to-hi-to
	311465	A1738660	Hs.208132	ESTs	to-hi-to
20	312073	AA682393	*Hs.119237	ESTs	to-hi-to
	312105	T81819	Hs.302251	ESTs	to-hi-to
	312108	T82331	Hs.127453	ESTs	to-hi-to
	312292	AW459103	Hs.151124	ESTs	to-hi-to
25	312313	AW293341	Hs.122505	ESTs, Weakly similar to t58022 hypothetical	to-hi-to
	312600	AW970985	Hs.230853	ESTs	to-hi-to
	312600	A1248774	Hs.127007	hypothetical protein FLJ11457	to-hi-to
	312621	A4869525	Hs.269860	ESTs	to-hi-to
	315097	A1676164	Hs.204339	ESTs	to-hi-to
30	313166	A1801098	Hs.151500	ESTs	to-hi-to
	313179	AA527670	Hs.131704	ESTs	to-hi-to
	313280	AW960454	Hs.222630	ESTs	to-hi-to
	315869	A1008910	Hs.152386	ESTs	to-hi-to
	314146	A1827237	Hs.282884	ESTs	to-hi-to
35	314305	A1280112	Hs.125232	Homo sapiens cDNA FLJ13266 fls, clone OV	to-hi-to
	314450	A1807931	Hs.104555	ESTs	to-hi-to
	314456	AA502917	Hs.150474	ESTs	to-hi-to
	314681	A1050587	Hs.152299	ESTs, Moderately similar to ALU5_HUMAN A	to-hi-to
	314916	AA548906	Hs.122244	ESTs	to-hi-to
40	315043	AA480658	Hs.130732	KIAA1575 protein	to-hi-to
	315074	AA526284	Hs.156729	Homo sapiens cDNA: FLJ21346 fls, clone C	to-hi-to
	315214	A1815527	Hs.34771	ESTs	to-hi-to
	315344	AW292176	Hs.245834	ESTs	to-hi-to
	315353	A1373949	Hs.279610	hypothetical protein FLJ10483	to-hi-to
45	315439	T78413	Hs.253636	ESTs	to-hi-to
	315628	R37257	Hs.184780	ESTs	to-hi-to
	315720	AA262988	Hs.163900	ESTs	to-hi-to
	315772	AW515373	Hs.271249	Homo sapiens cDNA FLJ13560 fls, clone PL	to-hi-to
	316841	AW136397	Hs.247572	ESTs	to-hi-to
50	316842	A1469680	Hs.170598	ESTs	to-hi-to
	316244	A1460781	Hs.224988	ESTs	to-hi-to
	316345	AW139408	Hs.152940	ESTs	to-hi-to
	316625	BE540090	Hs.122156	ESTs	to-hi-to
	316738	A4869085	Hs.123468	ESTs	to-hi-to
55	316868	A1660998	Hs.195002	ESTs	to-hi-to
	316905	AW138241	Hs.210846	ESTs	to-hi-to
	317224	X73618	*Hs.13029	sparganobionectin, ovcw and kazal-like d	to-hi-to
	317275	A1806444	Hs.202108	ESTs	to-hi-to
	317404	A1806867	Hs.126594	ESTs	to-hi-to
60	317488	AW071851	Hs.130628	ESTs	to-hi-to
	317916	A1565071	Hs.159983	ESTs	to-hi-to
	317939	A1861608	Hs.244760	ESTs	to-hi-to
	318480	T23814	Hs.133239	1-NIB Homo sapiens cDNA clone	to-hi-to
	318997	N46574	Hs.43838	ESTs	to-hi-to
65	320654	A1160015	Hs.118112	ESTs	to-hi-to
	320697	M6237	Hs.259108	ESTs	to-hi-to
	320787	AW186863	Hs.246240	ESTs	to-hi-to
	321023	AW294316	Hs.125608	ESTs	to-hi-to
	321899	AW972832	Hs.29468	ESTs	to-hi-to
70	322639	AA101897	Hs.211270	ESTs	to-hi-to
	322945	AA118990	Hs.186836	ESTs	to-hi-to
	323091	A1802456	Hs.210761	ESTs	to-hi-to
	323262	AL133990	Hs.190642	ESTs	to-hi-to
	323410	AW118683	Hs.154190	ESTs	to-hi-to
75	323945	AW145014	Hs.197746	ESTs	to-hi-to
	324098	AW172227	Hs.163686	Homo sapiens cDNA: FLJ22755 fls, clone K	to-hi-to
	324096	T78413	Hs.283696	ESTs	to-hi-to
	324674	AA541323	Hs.115831	ESTs	to-hi-to
	324713	A059330	Hs.313468	ESTs	to-hi-to
80	324790	A1334367	Hs.155337	ESTs	to-hi-to
	324804	A1892552	cdwd73812.x1 NCL_GCAP_1u24 Homo sapiens	to-hi-to	
	330728	A1805520	Hs.29672	ESTs	to-hi-to
	330760	H14588	Hs.30469	ESTs	to-hi-to

WO 02/098358

PCT/US02/17594

330776	AW953805	Hs.21887	ESTs	h-h-h
330824	AB037732	Hs.51441	KIAA1311 protein	h-h-h
331028	AK056552	Hs.28338	KIAA1546 protein	h-h-h
331046	N89593	Hs.191358	ESTs	h-h-h
331050	BE007967	Hs.155795	ESTs	h-h-h
331053	A949841	Hs.153466	ESTs, Moderately similar to ALU1_HUMAN A	h-h-h
331180	R44892	Hs.6540	Human DNA sequence from PAC 79N13 on chr	h-h-h
331313	AA761064	*Hs.80518	hypothetical protein	h-h-h
331337	N74392	Hs.50465	ESTs	h-h-h
331393	AW976438	Hs.17429	RFP-like protein	h-h-h
331432	A-923461	Hs.2345	ESTs	h-h-h
331517	AA765903	Hs.190877	H3 histone, family 3E (H3.3E)	h-h-h
331686	AW474890	Hs.182258	ESTs	h-h-h
332002	A87909	Hs.105104	ESTs	h-h-h
332043	AA371307	Hs.155055	ESTs	h-h-h
332295	AW770320	Hs.222413	ESTs	h-h-h
332314	R41396	Hs.101774	hypothetical protein FLJ23045	h-h-h
331517	AB037789	Hs.253395	sens domain, transmembrane domain (TM),	h-h-h
331562	AA934799	Hs.135228	ESTs, Moderately similar to ALU1_HUMAN A	h-h-h
331598	A829639	Hs.116252	ESTs, Moderately similar to ALU1_HUMAN A	h-h-h
332149	AA59177	Hs.172769	ESTs, Moderately similar to ALU1_HUMAN A	h-h-h
106099	NM_012068	Hs.9754	activating transcription factor 5	h-h-h
105726	NM_012068	Hs.9754	activating transcription factor 5	h-h-h
118025	AB20719	Hs.154682	DnaI (Hsp40) homolog, subfamily A, membe	h-h-h
3314915	A873735	Hs.187748	ESTs, Weakly similar to ALU1_HUMAN ALU S	h-h-h
331598	AB741508	Hs.188763	ESTs, Weakly similar to ALU1_HUMAN ALU S	h-h-h
324302	AW972771	Hs.252471	ESTs, Weakly similar to ALU1_HUMAN ALU S	h-h-h
331341	BEA10442	Hs.25240	Homo sapiens cDNA FLJ13466.6, clone PL	h-h-h
113783	AL359358	Hs.7041	hypothetical protein DKFZp682228	h-h-h
331552	AB89208	Hs.17283	hypothetical protein FLJ10890	h-h-h
103989	AA315983	Hs.105484	Homo sapiens regenerating gene type IV m	h-h-h
331492	AK021114	Hs.53913	hypothetical protein FLJ10252	h-h-h
110637	H03109	Hs.108920	HT018 protein	h-h-h
330814	AB95040	Hs.265398	ESTs, Weakly similar to transmembrane-r	h-h-h
331226	AA315703	Hs.199993	ESTs	h-h-h
102034	AB82474	Hs.230	fibronectin	h-h-h
134671	BE023265	Hs.302749	PKS6-binding protein 9 (63 kD)	h-h-h
331083	Y09763	Hs.22785	gamma-aminobutyric acid (GABA) A recepto	h-h-h
309575	AW188096	Hs.159476	glyceraldehyde-3-phosphate dehydrogenase	h-h-h
143432	CB6982	Hs.81876	growth factor receptor bound protein 10	h-h-h
132804	NM_005518	Hs.19689	3-hydroxy-3-methylglutaryl-Coenzyme A sy	h-h-h
302910	N77975	Hs.251577	hemoglobin, alpha 1	h-h-h
133731	N71725	*Hs.272572	hemoglobin, alpha 2	h-h-h
303297	AF070523	Hs.13423	Homo sapiens clone 24468 mRNA sequence	h-h-h
109732	AA328888	Hs.107476	ATP synthase, H+ transporting, mitochond	h-h-h
109731	AA328888	Hs.107476	ATP synthase, H+ transporting, mitochond	h-h-h
302123	AB013462	Hs.144931	ATPase, aminophospholipid transporter (A	h-h-h
131814	AB002438	Hs.28956	Homo sapiens mRNA from chromosome 5q21-2	h-h-h
104853	N94129	Hs.12959	hypothetical protein	h-h-h
302235	AL046987	Hs.163681	Homo sapiens mRNA; cDNA DKFZp664F112.1	h-h-h
300574	AL046943	Hs.161283	Homo sapiens mRNA; cDNA DKFZp669G2020.1	h-h-h
324678	AB90739	Hs.77868	ORF	h-h-h
331022	H03109	Hs.108920	HT018 protein	h-h-h
330430	H23350	Hs.21146	hypothetical protein FLJ23499	h-h-h
330601	U90918	Hs.82845	Homo sapiens cDNA: FLJ21930.6, clone H	h-h-h
101888	AF221521	Hs.8068	hematopoietic PEX-interacting protein	h-h-h
102959	AL036058	*Hs.75807	major histocompatibility complex, class	h-h-h
101363	M11521			h-h-h
133068	AA359886	Hs.232068	transcription factor 8 (represses Interl	h-h-h
332530	M31669	Hs.1735	Inhibin, beta E (active AB beta polypep	h-h-h
331777	NM_014785	Hs.47313	KIAA0298 gene product	h-h-h
100452	CB7742	Hs.241552	KIAA0235 protein	h-h-h
112989	NM_014867	Hs.5533	KIAA0711 gene product	h-h-h
320848	AB020691	Hs.198232	KIAA0684 protein	h-h-h
105182	AL133033	*Hs.4084	KIAA1025 protein	h-h-h
133905	AB028874	Hs.131476	KIAA1051 protein	h-h-h
331408	BE178953	Hs.23640	KIAA1105 protein	h-h-h
321441	AF107453	Hs.118458	Homo sapiens LUCA-15 protein mRNA, splic	h-h-h
131913	AW207440	Hs.185973	degenerative spermatocyte (homolog Dros	h-h-h
135424	U67611		transaldolase 1	h-h-h
128509	L4204	Hs.100724	peroxisome proliferative activated recep	h-h-h
332006	A150740	Hs.6241	phosphoinositide-3-kinase, regulatory su	h-h-h
311251	AB55652	Hs.197698	ESTs	h-h-h
314171	AB21895	Hs.193481	ESTs	h-h-h
105096	AW379378	Hs.170121	protein tyrosine phosphatase, receptor t	h-h-h
133740	AW52919	*Hs.170150	RAS, member RAS oncogene family-like	h-h-h
119521	W38038			h-h-h
119546	W38169			h-h-h
119559	W38197			h-h-h
133797	AL133821	Hs.76272	collagenase-binding protein 2	h-h-h
303096	AA942964	Hs.153593	ribosomal protein L18a	h-h-h
120255	AA169501	Hs.38710	hypothetical protein	h-h-h

322919	AA178955	Hs.271439	ESTs
300566	R34926		son of sevenless (Drosophila) homolog 1
302094	AT141617	Hs.108447	spinnoribcellular atrophy 4 [oligonucleotide
300914	AL120259	Hs.79591	transferrin growth factor, beta receptor
5	A037534	Hs.79059	thrombospondin 1
134565	AT570878	Hs.87409	UCP homologues 2 family, polypeptide
130117	U09641	Hs.150207	g-protein 17/18/19/20/21/22/23/24/25/26/27/28/29/30/31/32/33/34/35/36/37/38/39/40/41/42/43/44/45/46/47/48/49/50/51/52/53/54/55/56/57/58/59/60/61/62/63/64/65/66/67/68/69/70/71/72/73/74/75/76/77/78/79/80/81/82/83/84/85/86/87/88/89/90/91/92/93/94/95/96/97/98/99/100/101/102/103/104/105/106/107/108/109/110/111/112/113/114/115/116/117/118/119/120/121/122/123/124/125/126/127/128/129/130/131/132/133/134/135/136/137/138/139/140/141/142/143/144/145/146/147/148/149/150/151/152/153/154/155/156/157/158/159/160/161/162/163/164/165/166/167/168/169/170/171/172/173/174/175/176/177/178/179/180/181/182/183/184/185/186/187/188/189/190/191/192/193/194/195/196/197/198/199/200/201/202/203/204/205/206/207/208/209/210/211/212/213/214/215/216/217/218/219/220/221/222/223/224/225/226/227/228/229/230/231/232/233/234/235/236/237/238/239/240/241/242/243/244/245/246/247/248/249/250/251/252/253/254/255/256/257/258/259/260/261/262/263/264/265/266/267/268/269/270/271/272/273/274/275/276/277/278/279/280/281/282/283/284/285/286/287/288/289/290/291/292/293/294/295/296/297/298/299/300/301/302/303/304/305/306/307/308/309/310/311/312/313/314/315/316/317/318/319/320/321/322/323/324/325/326/327/328/329/330/331/332/333/334/335/336/337/338/339/340/341/342/343/344/345/346/347/348/349/350/351/352/353/354/355/356/357/358/359/360/361/362/363/364/365/366/367/368/369/370/371/372/373/374/375/376/377/378/379/380/381/382/383/384/385/386/387/388/389/390/391/392/393/394/395/396/397/398/399/400/401/402/403/404/405/406/407/408/409/410/411/412/413/414/415/416/417/418/419/420/421/422/423/424/425/426/427/428/429/430/431/432/433/434/435/436/437/438/439/440/441/442/443/444/445/446/447/448/449/450/451/452/453/454/455/456/457/458/459/460/461/462/463/464/465/466/467/468/469/470/471/472/473/474/475/476/477/478/479/480/481/482/483/484/485/486/487/488/489/490/491/492/493/494/495/496/497/498/499/500/501/502/503/504/505/506/507/508/509/510/511/512/513/514/515/516/517/518/519/520/521/522/523/524/525/526/527/528/529/530/531/532/533/534/535/536/537/538/539/540/541/542/543/544/545/546/547/548/549/550/551/552/553/554/555/556/557/558/559/560/561/562/563/564/565/566/567/568/569/570/571/572/573/574/575/576/577/578/579/580/581/582/583/584/585/586/587/588/589/590/591/592/593/594/595/596/597/598/599/600/601/602/603/604/605/606/607/608/609/610/611/612/613/614/615/616/617/618/619/620/621/622/623/624/625/626/627/628/629/630/631/632/633/634/635/636/637/638/639/640/641/642/643/644/645/646/647/648/649/650/651/652/653/654/655/656/657/658/659/660/661/662/663/664/665/666/667/668/669/670/671/672/673/674/675/676/677/678/679/680/681/682/683/684/685/686/687/688/689/690/691/692/693/694/695/696/697/698/699/700/701/702/703/704/705/706/707/708/709/710/711/712/713/714/715/716/717/718/719/720/721/722/723/724/725/726/727/728/729/730/731/732/733/734/735/736/737/738/739/740/741/742/743/744/745/746/747/748/749/750/751/752/753/754/755/756/757/758/759/760/761/762/763/764/765/766/767/768/769/770/771/772/773/774/775/776/777/778/779/780/781/782/783/784/785/786/787/788/789/790/791/792/793/794/795/796/797/798/799/800/801/802/803/804/805/806/807/808/809/810/811/812/813/814/815/816/817/818/819/820/821/822/823/824/825/826/827/828/829/830/831/832/833/834/835/836/837/838/839/840/841/842/843/844/845/846/847/848/849/850/851/852/853/854/855/856/857/858/859/860/861/862/863/864/865/866/867/868/869/870/871/872/873/874/875/876/877/878/879/880/881/882/883/884/885/886/887/888/889/890/891/892/893/894/895/896/897/898/899/900/901/902/903/904/905/906/907/908/909/910/911/912/913/914/915/916/917/918/919/920/921/922/923/924/925/926/927/928/929/930/931/932/933/934/935/936/937/938/939/940/941/942/943/944/945/946/947/948/949/950/951/952/953/954/955/956/957/958/959/960/961/962/963/964/965/966/967/968/96



PCT/US02/17594

	103145	X6822	Hs.16569	myosin binding protein C, slow type	to-hi-to
	301015	AV555272	Hs.22052	novel Ras family protein	to-hi-to
	311013	AA2274810	"Hs.153	isoform protein L7	to-hi-to
5	132650	AA276185	Hs.38022	EBF1	to-hi-to
	132649	AV173954	Hs.11248	endothelium inhibitor, Kuzal type 1	to-hi-to
	130989	AV972512	Hs.20985	3k3-associated polypeptide, 30kD	to-hi-to
	130781	AC030403	Hs.196263	Slc-20 related kinase	to-hi-to
	130585	AV506780	Hs.155223	steroid oxidase 2	to-hi-to
10	132722	AA316181	Hs.16163	slc1 transmembrane epithelial antigen of stomach	to-hi-to
	132620	SV9681	"Hs.17782	cardiac, pulmonary-associated protein	to-hi-to
	129523	ML1321	Hs.274509	T cell receptor gamma constant 2	to-hi-to
	321415	BG521687	Hs.1537	transmembrane alpha 4 superfamily member	to-hi-to
	31851	AV952654	Hs.3337	transmembrane alpha superfamily member	to-hi-to
5	133444	ML6378	Hs.3763	vascular endothelium growth factor	to-hi-to
	132657	AV393251	"Hs.2947	v-405 FBJ murine osteosarcoma viral onc	to-hi-to
	131328	AV393250	"Hs.2947	v-405 FBJ murine osteosarcoma viral onc	to-hi-to
	140372	AV215680	Hs.12619	transmembrane protein FLJ22316	to-hi-to
	140394	AA139551	Hs.172129	Home sapiens CDNA FLJ21409 f1, clone C	to-hi-to
20	103739	AA115713	gribs-30002.1 rat Striatum neuroepithel	to-hi-to	
	103738	AA115712	gribs-30002.1 rat Striatum neuroepithel	to-hi-to	
	103804	AA115713	gribs-30002.1 rat Striatum neuroepithel	to-hi-to	



PCT/US02/17594

[illegible]















WO 02/098358

PCT/US02/17594

		DT2855 T4229 A422943 WD0525 F07531 R25204 AF060467 T1974 T4 R38244 AW120570 A316170 W44675 H08965 A431931 W87260 H08924 R18966 R62083 N52000 BE541059 A3047315 F07042 BE206512 R62247 A3339023 R10468 R221904 W3187 W49532 L413414 A2126031 A347693 F07521 A21260876 A135283 F06385 N35364 R16629 A1163504 A803498 A4284445 N99020 A210742 A051965
5		A025246 A105291 A693408 AW065334 BE541793 A336740 AW082327 A024060 BE46505 A542333 AW049837 A2543679 H03859 A080470 A0658 A032035 T15765 BE205232 M61022 A467786 H543578 A143089 A135531 W51791 N25319 A23081 A112725 A104663 AW062076 A076734 A4675343 A348765 A147237 W67835 A343888 A347781 R084133072 R1458430 A027623 A255571 AA708716 A023236 A020315 A484963 A35171 A102604112893 A1465304 T35932 A1025784 A026795 A062988 A047059 R10146 N72323 R38857 A360444 R03659 N26252 R42579 R16570 A191075 A028202 W1661 A567782 A116989 W87538 A341305 W4476 T6129 A10447 R03036 BE191108 W67860 A564367 R38805 A332025 R30726 BE17300 A172948 AA201308 F0384 F02463 R4328 F10371 N85140 A135175 AW07065 A4851158 F08844 A35389 A131859 A33065 AW056144 A135446 H07858 AW107143 A181331 A458099A A857472 A444100 AW15766 W87238 F36094 A3473314 F131295 A023470 BE30295 BE328314 W161284 A021916 A4915406 A223849 A287965 A087020 A122735 F03067 N20841 A0406040 A872351 A094049 W87766 AA26256 A356358 A021912 A1149419 A4880044 W49533 182126 182138 185381 AW07859 R20242 A4235540 A80818 A56 A82139 A1276814 AW615475 N81668 A641406 A353233 A3034229 A116712 F088138 A357382 AW779420 A4878919 A130202 A111668 A484969 A165146 R5917 AA02502 A0255513 A1179027 R4306 R43783 A1179148 A160654 T85498 A422303 W0708 R52121 A027000
10		A350079 1124633 A0464320 H08996 1127691 A821259 A456323 A1748966 A24662 123278 A194361 A710137 F10117 T30222 T02075 A1194370 A3478103 H09852 T40917 A18205 R36566 H03035 T28294 R46268 H08925 R6142 A166673 A193018 R06282 A0230747 A1194455 R06190 R06578 F01995 R06510 R06340 A051346 A0551336 A0858496 R57012 B1227300 W2385 A041587 A4851749 T6408
20	100999	16380_1 H03765 A0202909 A0707561 A056872 M81600 T17043 A047038 R35469 A003438 A118163 A3338 A30428 A263902 A296231 A295488 A314063 BE018838 A3064417 BE365738 AW023935 W50213 BE27684 A3115325 A5418906 BE277221 A1149601 A252476 AW067824 N1_00903_0394 A4302015 H39427 A467801 H194365 M7385 F0278521 H6447 A52367 A1080981 AW362314 A1132025 A0753679 AW062910 A094728 A082713 A077390 A081065 A081368 BE220097 A020262 BE22282 A08762 A354386 AW14921 A457494 BE29345 A424950 A346644 A635390 A130215 A073632 A4587006 A577392 A533973 A5A8187 A156246 A663214 A0216776 A000994 DE081764 A07406 N2403 A588893 A1948709 A025372 A106858 A1671961 A113231 A51535 T127417 A4774708 N25131 A5467619 A374733 A336032 M61965 A365134 A405759 A32129 A331217 A301948 W573001 W58020 A035227 A4532804 A302277 A1797527 A469435 H07386 A154205 AW00569 A1085380 A561057 A103882 A0905378 AW065679 A0505677 N27709 A494545 A531152 A514646 H04386 D302400 A0579311 AW579374 A185411 AW3740 A515430 AW367310 AW176734 D11877 D11593 D11902 D12724 D11882 H39746 BE588335 H2648 A344338 BE16816 A037141 BE22011 A534681 A924726 A114684 A4455538 A020329 A017131 A4095499 C06934 AW088447 AW719930 A1185918 A1725300 AW789082 A035227 A4532804 A302277 A1797527 A469435 H07386 A154205 AW00569 A1085380 A561057 A103882 A0905378 AW065679 A0505677 N27709 A494545 A531152 A514646 H04386 D302400 A0579311 AW579374 A185411 AW3740 A515430 AW367310 AW176734 D11877 D11593 D11902 D12724 D11882 H39746 BE588335 H2648 A344338 BE16816 A037141 BE22011 A534681 A924726 A114684 A4455538 A020329 A017131 A4095499 C06934 AW088447 AW719930 A1185918 A1725300 AW789082 A035227 A4532804 A302277 A1797527 A469435 H07386 A154205 AW00569 A1085380 A561057 A103882 A0905378 AW065679 A0505677 N27709 A494545 A531152 A514646 H04386 D302400 A0579311 AW579374 A185411 AW3740 A515430 AW367310 AW176734 D11877 D11593 D11902 D12724 D11882 H39746 BE588335 H2648 A344338 BE16816 A037141 BE22011 A534681 A924726 A114684 A4455538 A020329 A017131 A4095499 C06934 AW088447 AW719930 A1185918 A1725300 AW789082 A035227 A4532804 A302277 A1797527 A469435 H07386 A154205 AW00569 A1085380 A561057 A103882 A0905378 AW065679 A0505677 N27709 A494545 A531152 A514646 H04386 D302400 A0579311 AW579374 A185411 AW3740 A515430 AW367310 AW176734 D11877 D11593 D11902 D12724 D11882 H39746 BE588335 H2648 A344338 BE16816 A037141 BE22011 A534681 A924726 A114684 A4455538 A020329 A017131 A4095499 C06934 AW088447 AW719930 A1185918 A1725300 AW789082 A035227 A4532804 A302277 A1797527 A469435 H07386 A154205 AW00569 A1085380 A561057 A103882 A0905378 AW065679 A0505677 N27709 A494545 A531152 A514646 H04386 D302400 A0579311 AW579374 A185411 AW3740 A515430 AW367310 AW176734 D11877 D11593 D11902 D12724 D11882 H39746 BE588335 H2648 A344338 BE16816 A037141 BE22011 A534681 A924726 A114684 A4455538 A020329 A017131 A4095499 C06934 AW088447 AW719930 A1185918 A1725300 AW789082 A035227 A4532804 A302277 A1797527 A469435 H07386 A154205 AW00569 A1085380 A561057 A103882 A0905378 AW065679 A0505677 N27709 A494545 A531152 A514646 H04386 D302400 A0579311 AW579374 A185411 AW3740 A515430 AW367310 AW176734 D11877 D11593 D11902 D12724 D11882 H39746 BE588335 H2648 A344338 BE16816 A037141 BE22011 A534681 A924726 A114684 A4455538 A020329 A017131 A4095499 C06934 AW088447 AW719930 A1185918 A1725300 AW789082 A035227 A4532804 A302277 A1797527 A469435 H07386 A154205 AW00569 A1085380 A561057 A103882 A0905378 AW065679 A0505677 N27709 A494545 A531152 A514646 H04386 D302400 A0579311 AW579374 A185411 AW3740 A515430 AW367310 AW176734 D11877 D11593 D11902 D12724 D11882 H39746 BE588335 H2648 A344338 BE16816 A037141 BE22011 A534681 A924726 A114684 A4455538 A020329 A017131 A4095499 C06934 AW088447 AW719930 A1185918 A1725300 AW789082 A035227 A4532804 A302277 A1797527 A469435 H07386 A154205 AW00569 A1085380 A561057 A103882 A0905378 AW065679 A0505677 N27709 A494545 A531152 A514646 H04386 D302400 A0579311 AW579374 A185411 AW3740 A515430 AW367310 AW176734 D11877 D11593 D11902 D12724 D11882 H39746 BE588335 H2648 A344338 BE16816 A037141 BE22011 A534681 A924726 A114684 A4455538 A020329 A017131 A4095499 C06934 AW088447 AW719930 A1185918 A1725300 AW789082 A035227 A4532804 A302277 A1797527 A469435 H07386 A154205 AW00569 A1085380 A561057 A103882 A0905378 AW065679 A0505677 N27709 A494545 A531152 A514646 H04386 D302400 A0579311 AW579374 A185411 AW3740 A515430 AW367310 AW176734 D11877 D11593 D11902 D12724 D11882 H39746 BE588335 H2648 A344338 BE16816 A037141 BE22011 A534681 A924726 A114684 A4455538 A020329 A017131 A4095499 C06934 AW088447 AW719930 A1185918 A1725300 AW789082 A035227 A4532804 A302277 A1797527 A469435 H07386 A154205 AW00569 A1085380 A561057 A103882 A0905378 AW065679 A0505677 N27709 A494545 A531152 A514646 H04386 D302400 A0579311 AW579374 A185411 AW3740 A515430 AW367310 AW176734 D11877 D11593 D11902 D12724 D11882 H39746 BE588335 H2648 A344338 BE16816 A037141 BE22011 A534681 A924726 A114684 A4455538 A020329 A017131 A4095499 C06934 AW088447 AW719930 A1185918 A1725300 AW789082 A035227 A4532804 A302277 A1797527 A469435 H07386 A154205 AW00569 A1085380 A561057 A103882 A0905378 AW065679 A0505677 N27709 A494545 A531152 A514646 H04386 D302400 A0579311 AW579374 A185411 AW3740 A515430 AW367310 AW176734 D11877 D11593 D11902 D12724 D11882 H39746 BE588335 H2648 A344338 BE16816 A037141 BE22011 A534681 A924726 A114684 A4455538 A020329 A017131 A4095499 C06934 AW088447 AW719930 A1185918 A1725300 AW789082 A035227 A4532804 A302277 A1797527 A469435 H07386 A154205 AW00569 A1085380 A561057 A103882 A0905378 AW065679 A0505677 N27709 A494545 A531152 A514646 H04386 D302400 A0579311 AW579374 A185411 AW3740 A515430 AW367310 AW176734 D11877 D11593 D11902 D12724 D11882 H39746 BE588335 H2648 A344338 BE16816 A037141 BE22011 A534681 A924726 A114684 A4455538 A020329 A017131 A4095499 C06934 AW088447 AW719930 A1185918 A1725300 AW789082 A035227 A4532804 A302277 A1797527 A469435 H07386 A154205 AW00569 A1085380 A561057 A103882 A0905378 AW065679 A0505677 N27709 A494545 A531152 A514646 H04386 D302400 A0579311 AW579374 A185411 AW3740 A515430 AW367310 AW176734 D11877 D11593 D11902 D12724 D11882 H39746 BE588335 H2648 A344338 BE16816 A037141 BE22011 A534681 A924726 A114684 A4455538 A020329 A017131 A4095499 C06934 AW088447 AW719930 A1185918 A1725300 AW789082 A035227 A4532804 A302277 A1797527 A469435 H07386 A154205 AW00569 A1085380 A561057 A103882 A0905378 AW065679 A0505677 N27709 A494545 A531152 A514646 H04386 D302400 A0579311 AW579374 A185411 AW3740 A515430 AW367310 AW176734 D11877 D11593 D11902 D12724 D11882 H39746 BE588335 H2648 A344338 BE16816 A037141 BE22011 A534681 A924726 A114684 A4455538 A020329 A017131 A4095499 C06934 AW088447 AW719930 A1185918 A1725300 AW789082 A035227 A4532804 A302277 A1797527 A469435 H07386 A154205 AW00569 A1085380 A561057 A103882 A0905378 AW065679 A0505677 N27709 A494545 A531152 A514646 H04386 D302400 A0579311 AW579374 A185411 AW3740 A515430 AW367310 AW176734 D11877 D11593 D11902 D12724 D11882 H39746 BE588335 H2648 A344338 BE16816 A037141 BE22011 A534681 A924726 A114684 A4455538 A020329 A017131 A4095499 C06934 AW088447 AW719930 A1185918 A1725300 AW789082 A035227 A4532804 A302277 A1797527 A469435 H07386 A154205 AW00569 A1085380 A561057 A103882 A0905378 AW065679 A0505677 N27709 A494545 A531152 A514646 H04386 D302400 A0579311 AW579374 A185411 AW3740 A515430 AW367310 AW176734 D11877 D11593 D11902 D12724 D11882 H39746 BE588335 H2648 A344338 BE16816 A037141 BE22011 A534681 A924726 A114684 A4455538 A020329 A017131 A4095499 C06934 AW088447 AW719930 A1185918 A1725300 AW789082 A035227 A4532804 A302277 A1797527 A469435 H07386 A154205 AW00569 A1085380 A561057 A103882 A0905378 AW065679 A0505677 N27709 A494545 A531152 A514646 H04386 D302400 A0579311 AW579374 A185411 AW3740 A515430 AW367310 AW176734 D11877 D11593 D11902 D12724 D11882 H39746 BE588335 H2648 A344338 BE16816 A037141 BE22011 A534681 A924726 A114684 A4455538 A020329 A017131 A4095499 C06934 AW088447 AW719930 A1185918 A1725300 AW789082 A035227 A4532804 A302277 A1797527 A469435 H07386 A154205 AW00569 A1085380 A561057 A103882 A0905378 AW065679 A0505677 N27709 A494545 A531152 A514646 H04386 D302400 A0579311 AW579374 A185411 AW3740 A515430 AW367310 AW176734 D11877 D11593 D11902 D12724 D11882 H39746 BE588335 H2648 A344338 BE16816 A037141 BE22011 A534681 A924726 A114684 A4455538 A020329 A017131 A4095499 C06934 AW088447 AW719930 A1185918 A1725300 AW789082 A035227 A4532804 A302277 A1797527 A469435 H07386 A154205 AW00569 A1085380 A561057 A103882 A0905378 AW065679 A0505677 N27709 A494545 A531152 A514646 H04386 D302400 A0579311 AW579374 A185411 AW3740 A515430 AW367310 AW176734 D11877 D11593 D11902 D12724 D11882 H39746 BE588335 H2648 A344338 BE16816 A037141 BE22011 A534681 A924726 A114684 A4455538 A020329 A017131 A4095499 C06934 AW088447 AW719930 A1185918 A1725300 AW789082 A035227 A4532804 A302277 A1797527 A469435 H07386 A154205 AW00569 A1085380 A561057 A103882 A0905378 AW065679 A0505677 N27709 A494545 A531152 A514646 H04386 D302400 A0579311 AW579374 A185411 AW3740 A515430 AW367310 AW176734 D11877 D11593 D11902 D12724 D11882 H39746 BE588335 H2648 A344338 BE16816 A037141 BE22011 A534681 A924726 A114684 A4455538 A020329 A017131 A4095499 C06934 AW088447 AW719930 A1185918 A1725300 AW789082 A035227 A4532804 A302277 A1797527 A469435 H07386 A154205 AW00569 A1085380 A561057 A103882 A0905378 AW065679 A0505677 N27709 A494545 A531152 A514646 H04386 D302400 A0579311 AW579374 A185411 AW3740 A515430 AW367310 AW176734 D11877 D11593 D11902 D12724 D11882 H39746 BE588335 H2648 A344338 BE16816 A037141 BE22011 A534681 A924726 A114684 A4455538 A020329 A017131 A4095499 C06934 AW088447 AW719930 A1185918 A1725300 AW789082 A035227 A4532804 A302277 A1797527 A469435 H07386 A154205 AW00569 A1085380 A561057 A103882 A0905378 AW065679 A0505677 N27709 A494545 A531152 A514646 H04386 D302400 A0579311 AW579374 A185411 AW3740 A515430 AW367310 AW176734 D11877 D11593 D11902 D12724 D11882 H39746 BE588335 H2648 A344338 BE16816 A037141 BE22011 A534681 A924726 A114684 A4455538 A020329 A017131 A4095499 C06934 AW088447 AW719930 A1185918 A1725300 AW789082 A035227 A4532804 A302277 A1797527 A469435 H07386 A154205 AW00569 A1085380 A561057 A103882 A0905378 AW065679 A0505677 N27709 A494545 A531152 A514646 H04386 D302400 A0579311 AW579374 A185411 AW3740 A515430 AW367310 AW176734 D11877 D11593 D11902 D12724 D11882 H39746 BE588335 H2648 A344338 BE16816 A037141 BE22011 A534681 A924726 A114684 A4455538 A020329 A017131 A4095499 C06934 AW088447 AW719930 A1185918 A1725300 AW789082 A035227 A4532804 A302277 A1797527 A469435 H07386 A154205 AW00569 A1085380 A561057 A103882 A0905378 AW065679 A0505677 N27709 A494545 A531152 A514646 H04386 D302400 A0579311 AW579374 A185411 AW3740 A515430 AW367310 AW176734 D11877 D11593 D11902 D12724 D11882 H39746 BE588335 H2648 A344338 BE16816 A037141 BE22011 A534681 A924726 A114684 A4455538 A020329 A017131 A4095499 C06934 AW088447 AW719930 A1185918 A1725300 AW789082 A035227 A4532804 A302277 A1797527 A469435 H07386 A154205 AW00569 A1085380 A561057 A103882 A0905378 AW065679 A0505677 N27709 A494545 A531152 A514646 H04386 D302400 A0579311 AW579374 A185411 AW3740 A515430 AW367310 AW176734 D11877 D11593 D11902 D12724 D11882 H39746 BE588335 H2648 A344338 BE16816 A037141 BE22011 A534681 A924726 A114684 A4455538 A020329 A017131 A4095499 C06934 AW088447 AW719930 A1185918 A1725300 AW789082 A035227 A4532804 A302277 A1797527 A469435 H07386 A154205 AW00569 A1085380 A561057 A103882 A0905378 AW065679 A0505677 N27709 A494545 A531152 A514646 H04386 D302400 A0579311 AW579374 A185411 AW3740 A515430 AW367310 AW176734 D11877 D11593 D11902 D12724 D11882 H39746 BE588335 H2648 A344338 BE16816 A037141 BE22011 A534681 A924726 A114684 A4455538 A020329 A017131 A4095499 C06934 AW088447 AW719930 A1185918 A1725300 AW789082 A035227 A4532804 A302277 A1797527 A469435 H07386 A154205 AW00569 A1085380 A561057 A103882 A0905378 AW065679 A0505677 N27709 A494545 A531152 A514646 H04386 D302400 A0579311 AW579374 A185411 AW3740 A515430 AW367310 AW176734 D11877 D11593 D11902 D12724 D11882 H39746 BE588335 H2648 A344338 BE16816 A03

PCT/US02/17594

1

PCT/US02/17594

1

WO 02/098358

PCT/US02/17594

5	29912_1	AD001914 NM 002670 M04082 R17185 AA393835 AW373682 AW373655 BE079653 R34474 W03207 AA359557 R1146 H69430 W89805 AT878782 AW001696 AD02051 AA904901 A132144 AN36457 AA834993 AB23882 H69432 AD042000 A359562 A336198 A30681 A785807 R01224 AW135608 H48238 AA888372 T28225 AA549213 AV457898 AW560431 H69047 T87338 T91020 R11147 NE7035 C00383 AA355808 AA563593 AW001696 BE076633 AA336388 AW373667 AW373678 AW373633 AW361388 H48329 AW373735 AA339158 AA350271 AA339821 A1133158 AW465680 AW56122 AW066237 AW666238 AW665108 AA564701 W03232 R01337 H61633 A5192764 A80801 AW5805 787439 H67322 AW027286 H6584 BE3889
		M0656 AA368768 AA57814 BE182030 AW02286 AB02284 BE281154 A030505 AW750490 AW040522 AW373288 AW240625 AA942435 AW579777 AD044828 BE272563 AW371988 AL047965 AW088830 AW560009 T28849 R45461 AW382220 BE030131 F00245 BE273339 AD147303 BE168320 T39134 BE079632 AD044448 AW181796 BE168256 AW060854 AW370278 BE037060 AW370278 BE037111 AA370383 AL047702 BE389257 AW042457 T10386 BE039800 BE007395 AW003201 A477478 AW003595 AW043377 BE004719 AW950564 AA022667 T56890 D56120 T29235 A015687 BE182058 BE182056 AW080238 M06567 AA347236 AW061866 AW176446 AA304671 AW0653735 T61714 AA316698 AA44815 AA345322 AA03485 AA488005 W52016 W35480 N67402 D06268 W25540 Y52847 D87278 D58961 BE619182 AA315188 AA308836 A112474 W79182 AD08564 H42266 AA301631 H18082 AA4113785 BE26097 AW515197 AA317390 BE11869 BE51780 AW545666 F11023 AA18000 AW029185 AA37608 BE54387 BE510311 AA482516 AA001808 AW159873 A1801283 A1565496 F00242 AW400696 AW564002 A1876753 T57653 T68110 AW060290 AW062910 AW045932 AW045932 AA102452 A135005 A018390 AA575797 AA382220 AW044222 BE332575 AW381471 AW063423 AA006800 AW570567 AW510672 A250777 AA033510 AW025709 AW15200 A0521563 A067954 A1148656 AW568533 A173825 AA463327 A027895 AW375542 AA455496 A4733014 A591384 R73530 R00323 AA413108 AA592638 AA414889 AW375592 AW568191 A474775 AW025657 A025770 AW338117 A10269111 A059442 A242300 A475476 H318718 AA396621 A113196 A4130023 H37940 T51625 AB06573 AW063871 A1179730 AA305757 A258455 M33565 A212013 AA336155 AW699969 T97325 AA345349 T81762 AA717981 A285052 A051398 BE329186 BE358582 A682601 A189284 AA345490 T5447 A292949 AW39207A AW68791 D82607 T48574 AW752038 C00300
20	25130_1	J04088 NM 01087 A071747 A145424 A046047 A2116574 BE26748 BE08381 A410077 AW56918 AW575045 H17813 BE081283 A070403 AW043327 BE094223 A410924 A171482 AW015337 A4137616 AB2744 AW03286 A472833 A34041 A785834 AA184838 AW235336 AW172627 AA056288 BE048383 A7314240 W16899 A060298 A289433 A4833734 AW184842 A420000 A420000 AA562873 W07031 BE020607 AA650003 A743147 A1590075 AA048274 A1129353 A536339 AA030313 A0245956 AW664319 A028183 A337454 AA370780 BE050282 A704047 A035038 A221521 A0074314 AW07889 A03332 A386969 A180828 AW173892 A1127498 A1079244 A191815 H17814 A570401 AW157854 T10273 A020682 A430025 AW063309 AW063309 A020682 A430025 A020682 H09977 AA306247 A4362501 AA039339 F05421 A2124475 AA30321 BE3904 A0409912 A19091542 AA402216 AW170382 AW170382 A111113 R94438 H37126 H6346 AA040928 A0490051 T29025 AW561071 L67277 L47276 A137591 BE38156 W2462 A474528 AA558223 BE065691 H69303 H67622 AA040348 A4377653 AW569911 BE15787 AA437415 H455437 AW253710 A075594 AA384435 H097194 AW57165 A101780 AW565661 A474775 A101780 A101780 A101780 A101780 A101780 A101780 A101780 A101780 A101780 A026783 AA135535 A425980 A047124 A2367334 AW514610 H69467 A026783 AA127289 A013456 A088720 A387374 AA383731 A1074253 A559577 A215786 A251819 A1459227 AA060022 A025234 T811838 A4580782 A423819 A1A00450 A427220 A1765284 A152007 AA078810 AA050794 A4729280 AA06328 AW788817 N11785 A060596 A4055822 A028874 A1588190 H684515 AW568515 A471314 AA449580 AA38416 AW569191 AW063307 W094770 A102095 A336964 A053030 A475819 W094770 D03035 D25155 AW021790 BE150884 F01875 A474414 A1941815 A545869 A4230788
		AW1076 R02440 AA38372 N47453 R01320 BE050653 N23862 W67442 AW248115 AA186459 W56117 A47571 A037045 W22606 W32746 AA40055 R01064 H02041 R06492 R3629 N27687 R0872 A3040278 F12437 H8051 A1045058 T3592 W5036 AA374465 AW1141 W15986 AA320401 BE009177 T66969 BE20995 A074648 BE30481 F0525 D04532 W1478 W9373 A7838 BE04847 AW1717 W11670 A1359302 R0472 D58901 A021221 A231280 H00300 M05749 H16447 T38128 H00300 R0055 T35424 AW03342 A073860 A4713587 A073366 W13973 A423827 R22120 AW051921 W19465 N5627 A0719711 M04576 W069840 A4236640 AW030222 A07982020 N31962 A1127153 A174829 W13739 AA375222 A055145 W48704 W2799 AA046584 A07054 AW593225 A4339058 A223852 A402344 AW362179 AA300069 AW012974 W19817 BE302618 AA314768 W50345 A1923055 W51050 AW590943 A2425023 A1959142 A4115108 A335985 A185783 N3367 N2309 M4651 A424654 BE366533 AW33887 AW239180 W63574 BE450517 AW01570 D5840 AW797405 BE385217 A05431 M45439 AW071987 A0249112 A119780 A03899 A140319 A068134 R05717 A22894 A430271 A1702381 A3961121 A771853 A582737 A032006 BE48221 BE53706 A772636 BE180448 A478419 AW06557 A4119063 BE54374 A1194589 A108281 A1625 W17102 Z14113 A050806 AA362755 AA37348 A0283633 AA045856 N13225 A11424 A41537 AA043320 N02055 A4337576 D59121 A2169373 A021698 H6833 W12658 A112626 W060665 N2579 AW395028 AW059822 A138284 N2564 N256232 W0046 A038241 AW13889 AW074776 A113040 A417209 R0381 T9770 A443687 AW17478 AW59592 A1151010 A0408282 A219733 A111653 A171895 A4313230 W2308 W3381 A1102222 A4951679 A096320 R21192 A1029945 A11861 A159810 A371476 R82157 A171355 N43589 A048000 AW564902 A425412 N6480 H20205 A481838 A541029 N6344 A7227 A7147405 A1128943 A425586 A051787 AW06399 A064131 A1436622 N6602 F09448 R64158 A088813 R22063 R0467 A4113696 A4415833 AA26233 A304105 T0405 AW035679 A812282 R0812 BE06521 A025122 W0494 T0406 BE06521 A025122 W0494 T0406 BE06521 W11220 H0384 AW003031 A435897 A090813 H91280 H0384 A090813 H91280 H0384 A090813 H91280 H0384 A090813 H91280 H0384 T33300 A395695 N94651 A4200033 AW061813 BE085681 N2624 A120350 BE062367 A502223 A197518 A1868152 N6017 BE172304 A148026 F2302 A039327 F0004 A0434961 A0178161 F36942 A02800 A093456 A1829015 A4305050 BE003393 H4607 BE01801 A1131895 BE2358 A051787 A107556 W63525 A03575 BE030313 A050764 A0381247 A542470 A1177 H7493 AW112886 A030351 AW171384 AW02955 AW02188 AW04552 A429306 A451085 BE18352 AW098775 AW103467 R06878 T39950 A050764 A0381247 AW080231 AW518979 A461741 BE070561 A0683472 A0684190 R2050 A537347 A1141112 A139141 H08223 A020534 A430384 A5715734 A0411008 A147858 A4577367 A028158 AW06549 A074055 A050694 A057363 AA342329 T8654 A23014 A4132975 A100227 A10151 N10220 A02000 A436675 W52320 A081658 AW065695 R09491 K2011 A11282165 AW057934 AA032301 A074035 R01207 A115240 AW04540 AW07409 A441790 R0866 A0688 BE12044 A06888 A872179 A47180 A050515 H6719 A067803 A0612993 F10069 A088577 A068547 A4128374 A100313 F01797 F02706 A064796 A0405493 H6929 A0845573 BE31251 AW01853 C01277 A684016 A151033 A151033 A428471 A003832 A823707 A47910 BE11975 A1058089 A054621 A099436 AA36478 A06886 A459776 AW06130 A423314 A094455 A0692226 A084908 A106289 A104573 A104947 A117265 AW17265 AW371138 A1359182 A4521835 AW06180 A07409 A441790 R0866 A0688 BE12044 A06888 A872179 A47180 A050515 H6719 A187803 A0612993 F10069 A088577 A068547 A4128374 A100313 F01797 F02706 A064796 A0405493 H6929 A0845573 BE31251 AW01853 C01277 A684016 A151033 A151033 A428471 A003832 A823707 A47910 BE11975 A1058089 A054621 A099436 AA36478 A06886 A459776 AW06130 A423314 A094455 A0692226 A084908 A106289 A104573 A104947 A117265 AW17265 AW371138 A1359182 A4521835 AW06180 A07409 A441790 R0866 A0688 BE12044 A06888 A872179 A47180 A050515 H6719 A187803 A0612993 F10069 A088577 A068547 A4128374 A100313 F01797 F02706 A064796 A0405493 H6929 A0845573 BE31251 AW01853 C01277 A684016 A151033 A151033 A428471 A003832 A823707 A47910 BE11975 A1058089 A054621 A099436 AA36478 A06886 A459776 AW06130 A423314 A094455 A0692226 A084908 A106289 A104573 A104947 A117265 AW17265 AW371138 A1359182 A4521835 AW06180 A07409 A441790 R0866 A0688 BE12044 A06888 A872179 A47180 A050515 H6719 A187803 A0612993 F10069 A088577 A068547 A4128374 A100313 F01797 F02706 A064796 A0405493 H6929 A0845573 BE31251 AW01853 C01277 A684016 A151033 A151033 A428471 A003832 A823707 A47910 BE11975 A1058089 A054621 A099436 AA36478 A06886 A459776 AW06130 A423314 A094455 A0692226 A084908 A106289 A104573 A104947 A117265 AW17265 AW371138 A1359182 A4521835 AW06180 A07409 A441790 R0866 A0688 BE12044 A06888 A872179 A47180 A050515 H6719 A187803 A0612993 F10069 A088577 A068547 A4128374 A100313 F01797 F02706 A064796 A0405493 H6929 A0845573 BE31251 AW01853 C01277 A684016 A151033 A151033 A428471 A003832 A823707 A47910 BE11975 A1058089 A054621 A099436 AA36478 A06886 A459776 AW06130 A423314 A094455 A0692226 A084908 A106289 A104573 A104947 A117265 AW17265 AW371138 A1359182 A4521835 AW06180 A07409 A441790 R0866 A0688 BE12044 A06888 A872179 A47180 A050515 H6719 A187803 A0612993 F10069 A088577 A068547 A4128374 A100313 F01797 F02706 A064796 A0405493 H6929 A0845573 BE31251 AW01853 C01277 A684016 A151033 A151033 A428471 A003832 A823707 A47910 BE11975 A1058089 A054621 A099436 AA36478 A06886 A459776 AW06130 A423314 A094455 A0692226 A084908 A106289 A104573 A104947 A117265 AW17265 AW371138 A1359182 A4521835 AW06180 A07409 A441790 R0866 A0688 BE12044 A06888 A872179 A47180 A050515 H6719 A187803 A0612993 F10069 A088577 A068547 A4128374 A100313 F01797 F02706 A064796 A0405493 H6929 A0845573 BE31251 AW01853 C01277 A684016 A151033 A151033 A428471 A003832 A823707 A47910 BE11975 A1058089 A054621 A099436 AA36478 A06886 A459776 AW06130 A423314 A094455 A0692226 A084908 A106289 A104573 A104947 A117265 AW17265 AW371138 A1359182 A4521835 AW06180 A07409 A441790 R0866 A0688 BE12044 A06888 A872179 A47180 A050515 H6719 A187803 A0612993 F10069 A088577 A068547 A4128374 A100313 F01797 F02706 A064796 A0405493 H6929 A0845573 BE31251 AW01853 C01277 A684016 A151033 A151033 A428471 A003832 A823707 A47910 BE11975 A1058089 A054621 A099436 AA36478 A06886 A459776 AW06130 A423314 A094455 A0692226 A084908 A106289 A104573 A104947 A117265 AW17265 AW371138 A1359182 A4521835 AW06180 A07409 A441790 R0866 A0688 BE12044 A06888 A872179 A47180 A050515 H6719 A187803 A0612993 F10069 A088577 A068547 A4128374 A100313 F01797 F02706 A064796 A0405493 H6929 A0845573 BE31251 AW01853 C01277 A684016 A151033 A151033 A428471 A003832 A823707 A47910 BE11975 A1058089 A054621 A099436 AA36478 A06886 A459776 AW06130 A423314 A094455 A0692226 A084908 A106289 A104573 A104947 A117265 AW17265 AW371138 A1359182 A4521835 AW06180 A07409 A441790 R0866 A0688 BE12044 A06888 A872179 A47180 A050515 H6719 A187803 A0612993 F10069 A088577 A068547 A4128374 A100313 F01797 F02706 A064796 A0405493 H6929 A0845573 BE31251 AW01853 C01277 A684016 A151033 A151033 A428471 A003832 A823707 A47910 BE11975 A1058089 A054621 A099436 AA36478 A06886 A459776 AW06130 A423314 A094455 A0692226 A084908 A106289 A104573 A104947 A117265 AW17265 AW371138 A1359182 A4521835 AW06180 A07409 A441790 R0866 A0688 BE12044 A06888 A872179 A47180 A050515 H6719 A187803 A0612993 F10069 A088577 A068547 A4128374 A100313 F01797 F02706 A064796 A0405493 H6929 A0845573 BE31251 AW01853 C01277 A684016 A151033 A151033 A428471 A003832 A823707 A47910 BE11975 A1058089 A054621 A099436 AA36478 A06886 A459776 AW06130 A423314 A094455 A0692226 A084908 A106289 A104573 A104947 A117265 AW17265 AW371138 A1359182 A4521835 AW06180 A07409 A441790 R0866 A0688 BE12044 A06888 A872179 A47180 A050515 H6719 A187803 A0612993 F10069 A088577 A068547 A4128374 A100313 F01797 F02706 A064796 A0405493 H6929 A0845573 BE31251 AW01853 C01277 A684016 A151033 A151033 A428471 A003832 A823707 A47910 BE11975 A1058089 A054621 A099436 AA36478 A06886 A459776 AW06130 A423314 A094455 A0692226 A084908 A106289 A104573 A104947 A117265 AW17265 AW371138 A1359182 A4521835 AW06180 A07409 A441790 R0866 A0688 BE12044 A06888 A872179 A47180 A050515 H6719 A187803 A0612993 F10069 A088577 A068547 A4128374 A100313 F01797 F02706 A064796 A0405493 H6929 A0845573 BE31251 AW01853 C01277 A684016 A151033 A151033 A428471 A003832 A823707 A47910 BE11975 A1058089 A054621 A099436 AA36478 A06886 A459776 AW06130 A423314 A094455 A0692226 A084908 A106289 A104573 A104947 A117265 AW17265 AW371138 A1359182 A4521835 AW06180 A07409 A441790 R0866 A0688 BE12044 A06888 A872179 A47180 A050515 H6719 A187803 A0612993 F10069 A088577 A068547 A4128374 A100313 F01797 F02706 A064796 A0405493 H6929 A0845573 BE31251 AW01853 C01277 A684016 A151033 A151033 A428471 A003832 A823707 A47910 BE11975 A1058089 A054621 A099436 AA36478 A06886 A459776 AW06130 A423314 A094455 A0692226 A084908 A106289 A104573 A104947 A117265 AW17265 AW371138 A1359182 A4521835 AW06180 A07409 A441790 R0866 A0688 BE12044 A06888 A872179 A47180 A050515 H6719 A187803 A0612993 F10069 A088577 A068547 A4128374 A100313 F01797 F02706 A064796 A0405493 H6929 A0845573 BE31251 AW01853 C01277 A684016 A151033 A151033 A428471 A003832 A823707 A47910 BE11975 A1058089 A054621 A099436 AA36478 A06886 A459776 AW06130 A423314 A094455 A0692226 A084908 A106289 A104573 A104947 A117265 AW17265 AW371138 A1359182 A4521835 AW06180 A07409 A441790 R0866 A0688 BE12044 A06888 A872179 A47180 A050515 H6719 A187803 A0612993 F10069 A088577 A068547 A4128374 A100313 F01797 F02706 A064796 A0405493 H6929 A0845573 BE31251 AW01853 C01277 A684016 A151033 A151033 A428471 A003832 A823707 A47910 BE11975 A1058089 A054621 A099436 AA36478 A06886 A459776 AW06130 A423314 A094455 A0692226 A084908 A106289 A104573 A104947 A117265 AW17265 AW371138 A1359182 A4521835 AW06180 A07409 A441790 R0866 A0688 BE12044 A06888 A872179 A47180 A050515 H6719 A187803 A0612993 F10069 A088577 A068547 A4128374 A100313 F01797 F02706 A064796 A0405493 H6929 A0845573 BE31251 AW01853 C01277 A684016 A151033 A151033 A428471 A003832 A823707 A47910 BE11975 A1058089 A054621 A099436 AA36478 A06886 A459776 AW06130 A423314 A094455 A0692226 A084908 A106289 A104573 A104947 A117265 AW17265 AW371138 A1359182 A4521835 AW06180 A07409 A441790 R0866 A0688 BE12044 A06888 A872179 A47180 A050515 H6719 A187803 A0612993 F10069 A088577 A068547 A4128374 A100313 F01797 F02706 A064796 A0405493 H6929 A0845573 BE31251 AW01853 C01277 A684016 A151033 A151033 A428471 A003832 A823707 A47910 BE11975 A1058089 A054621 A099436 AA36478 A06886 A459776 AW06130 A423314 A094455 A069



		BE530022 N65013 A621586 A146737 D19671 A1W152600 N54263 H73339 A9A10589 BE27344 BE560082 AW569012 AA313562 AW750034 BE072537 BE267491 AW732381 AA48336 D26674
102498	28114_1	USF42133 NM.013355 U48765 AW972510 AF012434 A1W162262 AE180579 BE467297 Y1151 AW680796 N3651 A003186 A6042840 A04134 A236360 A654522 A272957 BE553111 H10624 A314623 AW957178 AW749699 A9V4359 A818320 A81611 BE E383008 H81615 A61524 A505433 A68762 A674862 A6411 A1W65123 A1W65120 A1A3562 R27618 BE162204 A235645 R27441 A1950165 A00025 A65655 A656421 A740254 A A687591 A655195 A656782 H3767 A624130 A558336 A42524 N6421 A778106 A580225 A4315056 A6505428 A641740 A616226 A487584 A806488 A817456 N2537 A348564 A4873960 A173676 A2471427 A1141160 A21510 A1743909 A935171 A857553 A5W06215 R22665 A478688 A849777 A597824 A9A13449 A434906 A356202 A135915 H07915 A334505 A65749 A4351667 A206691 A345966 A461649 A4722613 A558001 A470955 R727113 A1313774 A7703815 A212652 A658204 A548485 A107761 A1A29270 A456365 A552348 A707765 A57876 A473767 A435296 A4229626 A861352 A868169 A1200207 W42762 A657630 R25067 A4548143 A604388 AW227276 N55172 R54783 R70265 R23061 R24742 A5646146 R25911 T56208 T52223 A561666 A2474611
131913	29330_1	AW70740 BE560098 BE22636 BE46791 BE50272 A7700391 BE348491 BE464713 BE46498 A10_00375 AF002695 E362402 AW723002 A357894 A1468139 A5268163 AW729144 A1A30032 A445678 AW88367 A448501 A438002 A4427347 BE162891 A439354 A5404320 T47547 AW80165 A552909 A6023967 C16061 A1A18006 N2306 A100246 A585788 A769126 A1270896 A581200 AW8161576 AW881631 BE170014 A0039929 R19395 N22526 A4809657
101840	2894_1	A223521 BE239171 AF05360 BE248325 AW07563 AW19331 M59305 A4317326 A4285297 BE519077 A4459462 A4828086 AW687040 AW357429 A5675813 A4361536 AW73811 A2828496 BE181033 BE177198 T60403 A4223079 T2677 A1W61546 N26286 A5165927 E5154541 A811545 A861116 A5617585 BE060638 AW064520 A685584 AW583009 A5527967 A1W70506 A1690287 A552476 A779184 AW87982 A81158740 AW70706 A6161260 A5672601 A1W361042 A1W4945 A1W59794 A563698 A5599013 A4238225 B655132 A4272610 A1914163 A701599 A4147840 AW795777 A4627978 A897344 BE219618 A788049 A390046 A418928 A465676 A175248 A449842 A563755 A573469 A4673835 A486275 A4069670 A4837553 A621518 A70342106 A1W5009 A476077 A1W50081 A1W50325 A277451 H08152 A277447 A1W01456 A686801 R54454 A752416 A787441 A1 BE358228 A00752 A1W51532 A450476 A1W51695 AW853106 AW457918 A932436 BE241436
25	131937	A907735 A05637 BE11345 A1432665 A140962 A1524050 A421702 A082401 AW291216 A207744 A361578 D0684 A361496 H50007 A278075 A572298 A427588 A5719185 A759306 A584110 D60423 A19334
131986	50779_1	AW70740 BE560098 BE22636 BE46791 BE50272 A7700391 BE348491 BE464713 BE46498 A10_00375 AF002695 E362402 AW723002 A357894 A1468139 A5268163 AW729144 A1A30032 A445678 AW88367 A448501 A438002 A4427347 BE162891 A581200 AW8161576 AW881631 BE170014 A0039929 R19395 N22526 A4809657
30	13682_1	AW076743 C03819 C03241 C11733 A4219339 BE80460 BE80509 L06846 Z32770 A04049112 A565139 A4234144 A207498 C0388 AW076743 C03819 C03241 C11733 A4219339 BE80460 BE80509 L06846 Z32770 A04049112 A565139 A4234144 A207498 C0388 AW076743 C03819 C03241 C11733 A4219339 BE80460 BE80509 L06846 Z32770 A04049112 A565139 A4234144 A207498 C0388
35	13682_1	AW076743 C03819 C03241 C11733 A4219339 BE80460 BE80509 L06846 Z32770 A04049112 A565139 A4234144 A207498 C0388 AW076743 C03819 C03241 C11733 A4219339 BE80460 BE80509 L06846 Z32770 A04049112 A565139 A4234144 A207498 C0388
101889	27210_13	AW076743 C03819 C03241 C11733 A4219339 BE80460 BE80509 L06846 Z32770 A04049112 A565139 A4234144 A207498 C0388 AW076743 C03819 C03241 C11733 A4219339 BE80460 BE80509 L06846 Z32770 A04049112 A565139 A4234144 A207498 C0388
131986	113870_1	AW076743 C03819 C03241 C11733 A4219339 BE80460 BE80509 L06846 Z32770 A04049112 A565139 A4234144 A207498 C0388 AW076743 C03819 C03241 C11733 A4219339 BE80460 BE80509 L06846 Z32770 A04049112 A565139 A4234144 A207498 C0388
10191	9877_1	AW076743 C03819 C03241 C11733 A4219339 BE80460 BE80509 L06846 Z32770 A04049112 A565139 A4234144 A207498 C0388 AW076743 C03819 C03241 C11733 A4219339 BE80460 BE80509 L06846 Z32770 A04049112 A565139 A4234144 A207498 C0388
103110	3166_1	AW076743 C03819 C03241 C11733 A4219339 BE80460 BE80509 L06846 Z32770 A04049112 A565139 A4234144 A207498 C0388 AW076743 C03819 C03241 C11733 A4219339 BE80460 BE80509 L06846 Z32770 A04049112 A565139 A4234144 A207498 C0388
45		AW076743 C03819 C03241 C11733 A4219339 BE80460 BE80509 L06846 Z32770 A04049112 A565139 A4234144 A207498 C0388 AW076743 C03819 C03241 C11733 A4219339 BE80460 BE80509 L06846 Z32770 A04049112 A565139 A4234144 A207498 C0388
103110	3166_1	AW076743 C03819 C03241 C11733 A4219339 BE80460 BE80509 L06846 Z32770 A04049112 A565139 A4234144 A207498 C0388 AW076743 C03819 C03241 C11733 A4219339 BE80460 BE80509 L06846 Z32770 A04049112 A565139 A4234144 A207498 C0388
50		AW076743 C03819 C03241 C11733 A4219339 BE80460 BE80509 L06846 Z32770 A04049112 A565139 A4234144 A207498 C0388 AW076743 C03819 C03241 C11733 A4219339 BE80460 BE80509 L06846 Z32770 A04049112 A565139 A4234144 A207498 C0388
103119	3394_1	AW076743 C03819 C03241 C11733 A4219339 BE80460 BE80509 L06846 Z32770 A04049112 A565139 A4234144 A207498 C0388 AW076743 C03819 C03241 C11733 A4219339 BE80460 BE80509 L06846 Z32770 A04049112 A565139 A4234144 A207498 C0388
65		AW076743 C03819 C03241 C11733 A4219339 BE80460 BE80509 L06846 Z32770 A04049112 A565139 A4234144 A207498 C0388 AW076743 C03819 C03241 C11733 A4219339 BE80460 BE80509 L06846 Z32770 A04049112 A565139 A4234144 A207498 C0388
103131	3166_1	AW076743 C03819 C03241 C11733 A4219339 BE80460 BE80509 L06846 Z32770 A04049112 A565139 A4234144 A207498 C0388 AW076743 C03819 C03241 C11733 A4219339 BE80460 BE80509 L06846 Z32770 A04049112 A565139 A4234144 A207498 C0388
75		AW076743 C03819 C03241 C11733 A4219339 BE80460 BE80509 L06846 Z32770 A04049112 A565139 A4234144 A207498 C0388 AW076743 C03819 C03241 C11733 A4219339 BE80460 BE80509 L06846 Z32770 A04049112 A565139 A4234144 A207498 C0388
103140	26990_2	AW076743 C03819 C03241 C11733 A4219339 BE80460 BE80509 L06846 Z32770 A04049112 A565139 A4234144 A207498 C0388 AW076743 C03819 C03241 C11733 A4219339 BE80460 BE80509 L06846 Z32770 A04049112 A565139 A4234144 A207498 C0388
132420	31972_1	AW076743 C03819 C03241 C11733 A4219339 BE80460 BE80509 L06846 Z32770 A04049112 A565139 A4234144 A207498 C0388 AW076743 C03819 C03241 C11733 A4219339 BE80460 BE80509 L06846 Z32770 A04049112 A565139 A4234144 A207498 C0388
80		AW076743 C03819 C03241 C11733 A4219339 BE80460 BE80509 L06846 Z32770 A04049112 A565139 A4234144 A207498 C0388 AW076743 C03819 C03241 C11733 A4219339 BE80460 BE80509 L06846 Z32770 A04049112 A565139 A4234144 A207498 C0388

5	103177	10888_1	DE243377 A3338310 X89141 A982390 AW140961 AA132168 AB163268 AA132950 AB16348 AB15502 AW40308 AA312277 AA311923
			AW163229 AW161244 DE5270 F07 120 A223128 BE323277 AA305886 AW247747 AA332597 AW550537 AW160538 AA311963
10	103180	10888_2	AA336171 NN_004462 L08106 L47700 R11849 AA423465 AW545655 AW163565 AW40308 AB16348 AB15502 AW40308 AA312277 AA311923
			AW163229 AW161244 DE5270 F07 120 A223128 BE323277 AA305886 AW247747 AA332597 AW550537 AW160538 AA311963
15	103181	10888_3	AA336171 NN_004462 L08106 L47700 R11849 AA423465 AW545655 AW163565 AW40308 AB16348 AB15502 AW40308 AA312277 AA311923
			AW163229 AW161244 DE5270 F07 120 A223128 BE323277 AA305886 AW247747 AA332597 AW550537 AW160538 AA311963
20	103182	10888_4	AA336171 NN_004462 L08106 L47700 R11849 AA423465 AW545655 AW163565 AW40308 AB16348 AB15502 AW40308 AA312277 AA311923
			AW163229 AW161244 DE5270 F07 120 A223128 BE323277 AA305886 AW247747 AA332597 AW550537 AW160538 AA311963
25	103183	10888_5	AA336171 NN_004462 L08106 L47700 R11849 AA423465 AW545655 AW163565 AW40308 AB16348 AB15502 AW40308 AA312277 AA311923
			AW163229 AW161244 DE5270 F07 120 A223128 BE323277 AA305886 AW247747 AA332597 AW550537 AW160538 AA311963
30	103184	10888_6	AA336171 NN_004462 L08106 L47700 R11849 AA423465 AW545655 AW163565 AW40308 AB16348 AB15502 AW40308 AA312277 AA311923
			AW163229 AW161244 DE5270 F07 120 A223128 BE323277 AA305886 AW247747 AA332597 AW550537 AW160538 AA311963
35	103185	10888_7	AA336171 NN_004462 L08106 L47700 R11849 AA423465 AW545655 AW163565 AW40308 AB16348 AB15502 AW40308 AA312277 AA311923
			AW163229 AW161244 DE5270 F07 120 A223128 BE323277 AA305886 AW247747 AA332597 AW550537 AW160538 AA311963
40	103186	10888_8	AA336171 NN_004462 L08106 L47700 R11849 AA423465 AW545655 AW163565 AW40308 AB16348 AB15502 AW40308 AA312277 AA311923
			AW163229 AW161244 DE5270 F07 120 A223128 BE323277 AA305886 AW247747 AA332597 AW550537 AW160538 AA311963
45	103187	10888_9	AA336171 NN_004462 L08106 L47700 R11849 AA423465 AW545655 AW163565 AW40308 AB16348 AB15502 AW40308 AA312277 AA311923
			AW163229 AW161244 DE5270 F07 120 A223128 BE323277 AA305886 AW247747 AA332597 AW550537 AW160538 AA311963
50	103188	10888_10	AA336171 NN_004462 L08106 L47700 R11849 AA423465 AW545655 AW163565 AW40308 AB16348 AB15502 AW40308 AA312277 AA311923
			AW163229 AW161244 DE5270 F07 120 A223128 BE323277 AA305886 AW247747 AA332597 AW550537 AW160538 AA311963
55	103189	10888_11	AA336171 NN_004462 L08106 L47700 R11849 AA423465 AW545655 AW163565 AW40308 AB16348 AB15502 AW40308 AA312277 AA311923
			AW163229 AW161244 DE5270 F07 120 A223128 BE323277 AA305886 AW247747 AA332597 AW550537 AW160538 AA311963
60	103190	10888_12	AA336171 NN_004462 L08106 L47700 R11849 AA423465 AW545655 AW163565 AW40308 AB16348 AB15502 AW40308 AA312277 AA311923
			AW163229 AW161244 DE5270 F07 120 A223128 BE323277 AA305886 AW247747 AA332597 AW550537 AW160538 AA311963
65	103191	10888_13	AA336171 NN_004462 L08106 L47700 R11849 AA423465 AW545655 AW163565 AW40308 AB16348 AB15502 AW40308 AA312277 AA311923
			AW163229 AW161244 DE5270 F07 120 A223128 BE323277 AA305886 AW247747 AA332597 AW550537 AW160538 AA311963
70	103192	10888_14	AA336171 NN_004462 L08106 L47700 R11849 AA423465 AW545655 AW163565 AW40308 AB16348 AB15502 AW40308 AA312277 AA311923
			AW163229 AW161244 DE5270 F07 120 A223128 BE323277 AA305886 AW247747 AA332597 AW550537 AW160538 AA311963
75	103193	10888_15	AA336171 NN_004462 L08106 L47700 R11849 AA423465 AW545655 AW163565 AW40308 AB16348 AB15502 AW40308 AA312277 AA311923
			AW163229 AW161244 DE5270 F07 120 A223128 BE323277 AA305886 AW247747 AA332597 AW550537 AW160538 AA311963
80	103194	10888_16	AA336171 NN_004462 L08106 L47700 R11849 AA423465 AW545655 AW163565 AW40308 AB16348 AB15502 AW40308 AA312277 AA311923
			AW163229 AW161244 DE5270 F07 120 A223128 BE323277 AA305886 AW247747 AA332597 AW550537 AW160538 AA311963

























WO 02/098358

PCT/US02/17594

5

127

2

5

134

20

25

40

15

55

103

5

70

5 100

30

134

128190	27120_14	AW040681 AW193490 AW973973 AW1949102 AW60284 AW64838 NM12896 AW54036 AW57334 AW542784 AW567405 AW411075
128191	27120_14	AW732563 AW538858 AW22913 AW1230069 AW67730 AW220064 AW631476 AW225236 AW72911 AW11564524 AW193184 AW228731 AW357255
104667	32675_1	AW230212 AW570090 AW228902 AW565053 AW229145 AW650240 AW653091 AW226143 AW579264 AW244482 AW185104 AW227671 AW533114
134790	1179_1	AW239923 BE46323 AW007234 AW126971 AW605333 AW009291 NM1082 AW664414 AW241833 AW62179 AW62386 BE250048
		AW20732 AW6291 AW6291 AW6291 AW6291 AW6291 AW6291 AW6291 AW6291 AW6291 AW6291 AW6291 AW6291 AW6291 AW6291 AW6291
		AW24464 AW5775 AW378943 AW42390 AW631313 AW781704 AW42895 AW149718 AW1018222 AW074666 AW03391 AW6363 AW59623
		H85342 AW3428137 AW340286 AW60555 AW615229 AW133116 AW2849 AW87642 AW340078 AW343380 AW501549 AW342052 AW6281
		AW172330 AW110693 AW063589 AW064745 AW26219 AW1049331 AW207309 AW149643 AW003343 AW181440 AW1495335 AW42965 AW22827 AW317313
		AW1133874 AW78803 AW0021265 AW1398965 AW545337 AW340231 AW149643 AW72165 AW1133874 AW1495335 AW42965 AW22827 AW317313
		AW665530 AW119077 AW1141996 AW74040 AW119077 AW119077 AW119077 AW119077 AW119077 AW119077 AW119077 AW119077 AW119077 AW119077 AW119077
		AW160135 AW119440 AW1495335 AW008312 AW369628 AW383327 AW6811 AW119440 AW1495335 AW008312 AW369628 AW383327 AW6811 AW119440 AW1495335
		AW219619 AW150246 AW119243 AW4272 AW68003 AW005693 AW6811 AW119440 AW1495335 AW008312 AW369628 AW383327 AW6811 AW119440 AW1495335
		AW225326 AW67076 AW4324 AW6811 AW4272 AW68003 AW005693 AW6811 AW119440 AW1495335 AW008312 AW369628 AW383327 AW6811 AW119440 AW1495335
		AW176345 AW176345 AW176345 AW176345 AW176345 AW176345 AW176345 AW176345 AW176345 AW176345 AW176345 AW176345 AW176345 AW176345 AW176345
		AW044667 AW04507 AW1191370 AW67098 AW725413 AW720257 AW00020 AW61572 AW6842 AW1631576 AW6811 AW119440 AW1495335 AW008312 AW369628
		AW0001 F11551 F140496 F2783 F36062 NM100361 AW36251 AW421696 F3338 F3339 F36189 F30349 AW44587 AW119440 AW1495335 AW008312 AW369628
		AW1193214 AW1170 AW330081 AW119655 AW107806 AW002747 AW030800 AW337970 AW6811 AW1495335 AW42965 AW22827 AW317313 AW317313
		AW525362 AW1495335 AW008312 AW117313 AW117313 AW117313 AW117313 AW117313 AW117313 AW117313 AW117313 AW117313 AW117313 AW117313 AW117313
		AW1761193 AW0001 AW67098 AW725413 AW720257 AW00020 AW61572 AW6842 AW1631576 AW6811 AW119440 AW1495335 AW008312 AW369628
		AW641147 AW117313 AW117313 AW117313 AW117313 AW117313 AW117313 AW117313 AW117313 AW117313 AW117313 AW117313 AW117313 AW117313 AW117313
120308	187025_2	AW323278 AW1341975 AW505011 AW591064 AW1196979
113168	8286_1	AW328102 AW11962379 AW4336 AW223996 AW001611 AW1190714 AW0084 AW0084 AW0084 AW0084 AW0084 AW0084 AW0084 AW0084 AW0084
		AW0084 AW0084 AW0084 AW0084 AW0084 AW0084 AW0084 AW0084 AW0084 AW0084 AW0084 AW0084 AW0084 AW0084 AW0084
		AW247829 AW0287 AW031 AW67220 AW045034 AW227955 AW353766 AW1997541 AW6140 AW66466 AW187235 BE30044 BE467445
		AW198945 AW333281 AW275429 AW340430 AW67692 AW008312 AW176543 AW467238 AW24147 AW251768 AW17345 AW33438 AW33438
		AW1294018 AW0012 AW026273 AW1495335 AW42965 AW198979 AW119440 AW1495335 AW42965 AW198979 AW119440 AW1495335 AW42965 AW198979
		AW034547 AW33305 F3212 AW08074 AW780493 AW0157 AW119440 AW1495335 AW42965 AW198979 AW119440 AW1495335 AW42965 AW198979
		BE013348 AW004405 R22239 AW005136 AW127004 AW07427 R22184 AW123608 AW049455 AW33568 R74220 AW077999 AW17300 AW17300
		AW198979 AW35686 AW436272 AW036878 AW689806 AW360561 R26449 AW86747 AW518513 AW20011 AW232796 AW00119 AW44772
		AW233854 AW026320 AW450294 AW0256
		XW6932 NM10004 AW0150 AW02303 AW0150 AW0150 AW0150 AW0150 AW0150 AW0150 AW0150 AW0150 AW0150 AW0150 AW0150 AW0150
		BE238524 AW036182 AW67599 AW78838 AW027306 BE072856 AW036176 AW73040 AW009694 BE047106 R03995 BE2358
		AW575008 AW343555 AW197935 AW232001 AW1495335 AW008312 AW176543 AW467238 AW24147 AW251768 AW17345 AW33438 AW33438
		AW1025432 R71001 AW225644 AW002949 AW64482 AW008312 AW133662 R71002 AW119440 AW17793 AW17793 AW17793 AW17793 AW17793 AW17793
		BE0256 AW183479 AW1495335 AW134730 AW373345 AW015055 AW119440 AW1495335 AW42965 AW198979 AW119440 AW1495335 AW42965 AW198979
		AW04532 BE04091 AW721256 AW236972 AW133622 AW78911 AW54515 AW68320 AW01021 AW119440 AW17793 AW17793 AW17793 AW17793
		AW22539 AW10637 R3980 AW87924 AW576472 AW02642 AW34366 AW1992204 AW06921 R03816 AW129376 AW02116 AW179637
		AW002644 AW519135 AW065649 AW007275 AW133662 AW42738
		AW278931 AW119440 AW1495335 AW42965 AW198979 AW119440 AW1495335 AW42965 AW198979 AW119440 AW1495335 AW42965 AW198979
		AW17793 AW17793 AW17793 AW17793 AW17793 AW17793 AW17793 AW17793 AW17793 AW17793 AW17793 AW17793 AW17793 AW17793
		AW161500 AW684379 AW162186 AW2278 AW166606 AW2278 AW166606 AW2278 AW166606 AW2278 AW166606 AW2278 AW166606 AW2278
		AW028127 AW0030 BE0220 F19093 F16783
		BE251623 BE251539 AW112200 AW007314 AW0472 AW22993 AW112200 AW112200 AW112200 AW112200 AW112200 AW112200 AW112200 AW112200
		AW0084 AW0084 AW0084 AW0084 AW0084 AW0084 AW0084 AW0084 AW0084 AW0084 AW0084 AW0084 AW0084 AW0084 AW0084
		AW014890 NM104287 AW014879 AW00306 AW243229 AW120699 AW030111 BE206236 AW307007 AW030111 AW030111 AW030111 AW030111 AW030111
		AW004874 AW000712 AW02044 AW002643 AW125782 AW8157 AW8267 AW290268 AW8880 AW366890 AW44807 AW184543 AW44807 AW184543
		AW004898 AW1114 AW8887 AW3596 AW131577 AW130045 AW002278 AW372329 AW00537 AW174535 AW44807 AW174535 AW44807 AW174535
		AW242157 AW8931 AW54535 AW54535 AW19009 AW19009 AW19009 AW19009 AW19009 AW19009 AW19009 AW19009 AW19009 AW19009
		AW028127 AW0030 BE0220 F19093 F16783
		BE251623 BE251539 AW112200 AW007314 AW0472 AW22993 AW112200 AW112200 AW112200 AW112200 AW112200 AW112200 AW112200 AW112200
		AW0084 AW0084 AW0084 AW0084 AW0084 AW0084 AW0084 AW0084 AW0084 AW0084 AW0084 AW0084 AW0084 AW0084 AW0084
		AW014890 NM104287 AW014879 AW00306 AW243229 AW120699 AW030111 BE206236 AW307007 AW030111 AW030111 AW030111 AW030111
		AW004874 AW000712 AW02044 AW002643 AW125782 AW8157 AW8267 AW290268 AW8880 AW366890 AW44807 AW184543 AW44807 AW184543
		AW004898 AW1114 AW8887 AW3596 AW131577 AW130045 AW002278 AW372329 AW00537 AW174535 AW44807 AW174535 AW44807 AW174535
		AW242157 AW8931 AW54535 AW54535 AW19009 AW19009 AW19009 AW19009 AW19009 AW19009 AW19009 AW19009 AW19009 AW19009
		AW028127 AW0030 BE0220 F19093 F16783
		BE251623 BE251539 AW112200 AW007314 AW0472 AW22993 AW112200 AW112200 AW112200 AW112200 AW112200 AW112200 AW112200 AW112200
		AW0084 AW0084 AW0084 AW0084 AW0084 AW0084 AW0084 AW0084 AW0084 AW0084 AW0084 AW0084 AW0084 AW0084 AW0084
		AW014890 NM104287 AW014879 AW00306 AW243229 AW120699 AW030111 BE206236 AW307007 AW030111 AW030111 AW030111 AW030111
		AW004874 AW000712 AW02044 AW002643 AW125782 AW8157 AW8267 AW290268 AW8880 AW366890 AW44807 AW184543 AW44807 AW184543
		AW004898 AW1114 AW8887 AW3596 AW131577 AW130045 AW002278 AW372329 AW00537 AW174535 AW44807 AW174535 AW44807 AW174535
		AW242157 AW8931 AW54535 AW54535 AW19009 AW19009 AW19009 AW19009 AW19009 AW19009 AW19009 AW19009 AW19009 AW19009
		AW028127 AW0030 BE0220 F19093 F16783
		BE251623 BE251539 AW112200 AW007314 AW0472 AW22993 AW112200 AW112200 AW112200 AW112200 AW112200 AW112200 AW112200 AW112200
		AW0084 AW0084 AW0084 AW0084 AW0084 AW0084 AW0084 AW0084 AW0084 AW0084 AW0084 AW0084 AW0084 AW0084 AW0084
		AW014890 NM104287 AW014879 AW00306 AW243229 AW120699 AW030111 BE206236 AW307007 AW030111 AW030111 AW030111 AW030111
		AW004874 AW000712 AW02044 AW002643 AW125782 AW8157 AW8267 AW290268 AW8880 AW366890 AW44807 AW184543 AW44807 AW184543
		AW004898 AW1114 AW8887 AW3596 AW131577 AW130045 AW002278 AW372329 AW00537 AW174535 AW44807 AW174535 AW44807 AW174535
		AW242157 AW8931 AW54535 AW54535 AW19009 AW19009 AW19009 AW19009 AW19009 AW19009 AW19009 AW19009 AW19009 AW19009
		AW028127 AW0030 BE0220 F19093 F16783
		BE251623 BE251539 AW112200 AW007314 AW0472 AW22993 AW112200 AW112200 AW112200 AW112200 AW112200 AW112200 AW112200 AW112200
		AW0084 AW0084 AW0084 AW0084 AW0084 AW0084 AW0084 AW0084 AW0084 AW0084 AW0084 AW0084 AW0084 AW0084 AW0084
		AW014890 NM104287 AW014879 AW00306 AW243229 AW120699 AW030111 BE206236 AW307007 AW030111 AW030111 AW030111 AW030111
		AW004874 AW000712 AW02044 AW002643 AW125782 AW8157 AW8267 AW290268 AW8880 AW366890 AW44807 AW184543 AW44807 AW184543
		AW004898 AW1114 AW8887 AW3596 AW131577 AW130045 AW002278 AW372329 AW00537 AW174535 AW44807 AW174535 AW44807 AW174535
		AW242157 AW8931 AW54535 AW54535 AW19009 AW19009 AW19009 AW19009 AW19009 AW19009 AW19009 AW19009 AW19009 AW19009
		AW028127 AW0030 BE0220 F19093 F16783
		BE251623 BE251539 AW112200 AW007314 AW0472 AW22993 AW112200 AW112200 AW112200 AW112200 AW112200 AW112200 AW112200 AW112200
		AW0084 AW0084 AW0084 AW0084 AW0084 AW0084 AW0084 AW0084 AW0084 AW0084 AW0084 AW0084 AW0084 AW0084 AW0084
		AW014890 NM104287 AW014879 AW00306 AW243229 AW120699 AW030111 BE206236 AW307007 AW030111 AW030111 AW030111 AW030111
		AW004874 AW000712 AW02044 AW002643 AW125782 AW8157 AW8267 AW290268 AW8880 AW366890 AW44807 AW184543 AW44807 AW184543
		AW004898 AW1114 AW8887 AW3596 AW131577 AW130045 AW002278 AW372329 AW00537 AW174535 AW44807 AW174535 AW44807 AW174535
		AW242157 AW8931 AW54535 AW54535 AW19009 AW19009 AW19009 AW19009 AW19009 AW19009 AW19009 AW19009 AW19009 AW19009
		AW028127 AW0030 BE0220 F19093 F16783
		BE251623 BE251539 AW112200 AW007314 AW0472 AW22993 AW112200 AW112200 AW112200 AW112200 AW112200 AW112200 AW112200 AW112200
		AW0084 AW0084 AW0084 AW0084 AW0084 AW0084 AW0084 AW0084 AW0084 AW0084 AW0084 AW0084 AW0084 AW0084 AW0084
		AW014890 NM104287 AW014879 AW00306 AW243229 AW120699 AW030111 BE206236 AW307007 AW030111 AW030111 AW030111 AW030111
		AW004874 AW000712 AW02044 AW002643 AW125782 AW8157 AW8267 AW290268 AW8880 AW366890 AW44807 AW184543 AW44807 AW184543
		AW004898 AW1114 AW8887 AW3596 AW131577 AW130045 AW002278 AW372329 AW00537 AW174535 AW44807 AW174535 AW44807 AW174535
		AW242157 AW8931 AW54535 AW54535 AW19009 AW19009 AW19009 AW19009 AW19009 AW19009 AW19009 AW19009 AW19009 AW19009
		AW028127 AW0030 BE0220 F19093 F16783
		BE251623 BE251539 AW112200 AW007314 AW0472 AW22993 AW112200 AW112200 AW112200 AW112200 AW112200 AW112200 AW112200 AW112200
		AW0084 AW0084 AW0084 AW0084 AW0084 AW0084 AW0084 AW0084 AW0084 AW0084 AW0084 AW0084 AW0084 AW0084 AW0084
		AW014890 NM104287 AW014879 AW00306 AW243229 AW120699 AW030111 BE206236 AW307007 AW030111 AW030111 AW030111 AW030111
		AW004874 AW000712 AW02044 AW002643 AW125782 AW8157 AW8267 AW290268 AW8880 AW366890 AW44807 AW184543 AW44807 AW184543
		AW004898 AW1114 AW8887 AW3596 AW131577 AW130045 AW002278 AW372329 AW00537 AW174535 AW44807 AW174535 AW44807 AW174535
		AW242157 AW8931 AW54535 AW54535 AW19009 AW19009 AW19009 AW19009 AW19009 AW19009 AW19009 AW19009 AW19009 AW19009
		AW028127 AW0030 BE0220 F19093 F16783
		BE251623 BE251539 AW112200 AW007314 AW0472 AW22993 AW112200 AW112200 AW112200 AW112200 AW112200 AW112200 AW112200 AW112200
		AW0084 AW0084 AW0084 AW0084 AW0084 AW0084 AW0084 AW0084 AW0084 AW0084 AW0084 AW0084 AW0084 AW0084 AW0084
		AW014890 NM104287 AW014879 AW00306 AW243229 AW120699 AW030111 BE206236 AW307007 AW030111 AW030111 AW030111 AW030111
		AW004874 AW000712 AW02044 AW002643 AW125782 AW8157 AW8267 AW290268 AW8880 AW366890 AW44807 AW184543 AW44807 AW184543
		AW004898 AW1114 AW8887 AW3596 AW131577 AW130045 AW002278 AW372329 AW00537 AW174535 AW44807 AW174535 AW44807 AW174535
		AW242157 AW8931 AW54535 AW54535 AW19009 AW19009 AW19009 AW19009 AW19009 AW19009 AW19009 AW19009 AW19009 AW19009
		AW028127 AW0030 BE0220 F19093 F16783
		BE251623 BE251539 AW112200 AW007314 AW0472 AW22993 AW112200 AW112200 AW112200 AW112200 AW112200 AW112200 AW112200 AW112200
		AW0084 AW0084 AW0084 AW0084 AW0084 AW0084 AW0084 AW0084 AW0084 AW0084 AW0084 AW0084 AW0084 AW0084 AW0084
		AW014890 NM104287 AW014879 AW00306 AW243229 AW120699 AW030111 BE206236 AW307007 AW030111 AW030111 AW030111 AW030111
		AW004874 AW000712 AW02044 AW002643 AW125782 AW8157 AW8267 AW290268 AW8880 AW366890 AW44807 AW184543 AW44807 AW184543
		AW004898 AW1114 AW8887 AW3596 AW131577 AW130045 AW002278 AW372329 AW00537 AW174535 AW44807 AW174535 AW44807 AW174535
		AW242157 AW8931 AW54535 AW54535 AW19009 AW19009 AW19009 AW19009 AW19009 AW19009 AW19009 AW19009 AW19009 AW19009
		AW028127 AW0030 BE0220 F19093 F16783
		BE251623 BE251539 AW112200 AW007314 AW0472 AW22993 AW112200 AW112200 AW112200 AW112200 AW112200 AW112200 AW112200 AW112200
		AW0084 AW0084 AW0084 AW0084 AW0084 AW0084 AW0084 AW0084 AW0084 AW0084 AW0084 AW0084 AW0084 AW0084 AW0084
		AW014890 NM104287 AW014879 AW00306 AW243229 AW120699 AW030111 BE206236 AW307007 AW030111 AW030111 AW030111 AW030111
		AW004874 AW000712 AW02044 AW002643 AW125782 AW8157 AW8267 AW290268 AW8880 AW366890 AW44807 AW184543 AW44807 AW184543
		AW004898 AW1114 AW8887 AW3596 AW131577 AW130045 AW002278 AW372329 AW00537 AW174535 AW44807 AW174535 AW44807 AW174535
		AW242157 AW8931 AW54535 AW54535 AW19009 AW19009 AW19009 AW19009 AW19009 AW19009 AW19009 AW19009 AW19009 AW19009
		AW028127 AW0030 BE0220 F19093 F16783
		BE251623 BE251539 AW112200 AW007314 AW0472 AW22993 AW112200 AW112200 AW112200 AW112200 AW112200 AW112200 AW112200 AW112200
		AW0084 AW0084 AW0084 AW0084 AW0084 AW0084 AW0084 AW0084 AW0084 AW0084 AW0084 AW0084 AW0084 AW0084 AW0084
		AW014890 NM104287 AW014879 AW00306 AW243

			T31825 N15599 W15471 H30034 AW69567 AB44311 N69003 D11866 AA04066 AW20826 AB62215 AA85454 RB5086 RH9957 N11259 AA42938 AC08526 D72680 AW162736 AA158136 AA01506 AA690420 AA197393 AS171336 AB193322 AE27967 AW16786 A119259 AA158136 AB67213 AW73294 AA125846 AW171452 AB69147 AA134528 AW23478 AB136041 AB301766 AW508201 AW213507 A133860 AW166321 AB72028 AW565445 AB92731 AB165896 AW69552 AB711178 AB14568 AW21452 AW20240 AW960151 AW237650 AW275444 AB78195	
5			AW502166 W27807 AC256985 AB86416 AB42423 AB741612 AB88244 D2649422 H6303 AW339154 AW25950 A000160 AW513178 A250166 AW2191 AA221988 AA259485 C01115 RB3652 AS30962 AA330630 AB846438 AW025325 W75947 AB706390 AB267672 W74079 A227654 A202272 AB242436 AC265485 NW0717 AA228927 AC202792 AB68213 AS157102 H44343 A112442 AW540651 AW51561 C01162 AC212346 AC080685 AA177392 W17708 AB6329 AB21930 AA045814 A1733456 U33046 A4189109 T31714 H110126 AA730493 F65971 R42427 AB313048 BE044060 A220005 A1283796 R39299 A2716749 AW14376 R46634 A305345 A226908 H02099 F10707 AA110649 H07336 H07345 AA125964 AA03438 AA047635 BE160887 AW053643 AC359126 BE393582 AW007442 AA235675 W06545 AA656404 H83331 BE274671 AA242039 A1168132 W26186 AB051820 AW507615 A2235937 A4232711 A140312 AW6924 A277475 A112610 AA74307 AB25066 AW16008 A113084 A226677 AB65332 AB21964 AA445595 AS157102 AS33081 RA6860 A13084 AB647535 AW676373 AB977525 W06824 AA173541 AA651611 A232200 AB277867 AA477767 A130284 AB706506 A114540 H2051 AB87059 AW73346 A242452 AS69069 AW24832 AB71519 AW002605 AB934054 A500194 AA876721 AW182176 AA25666 AS190769 A338251 AA153278 AS35679 A1199434 Z39891 AB805062 AA58275 BE250162 BE259656 NM_002439 U81581 AA421716 AA723916 N2236 H68382 A187671 AB36600 AW354468 BE250719 AB354498 AW03729 AB30089 AW173270 AC4278 BE11638 BE14324 AW06036 AB695663 J00140 AW575796 A431433 A1404147 R05447 A471432 T29009 W47330 AC24555 AS19077 A272820 AB457802 AB27163 A2321263 AS1952425 A472763 A170752 A0044683 BE467765 H23667 J00146 AC252992 A194511 A694127 A352165 AA250600 A0M0148 AB00605 A216565 A227746 A404147 A405130 A333571 A1871738 AW103565 AA010676 A2389782 A1131225 AA1700499 R91775 A4778381 AA455309 R00894 T9134 AA482938 AW511327 AC29077 A00040 BE03981 H25294 BE37236 AW011317 BE263372 J00149 J00605 V00567 BE263613 BE262099 BE26506 BE391734 BE256338 A0041848 J04610 A476934 AA489055 A128465 AA20816 AA280333 AC303332 A5572941 AW572255 A4725857 A114392 A514362 AB86480 AA63881 NM_000791 AA447680 AB360897 H49531 A4129464 BE071301 W23560 A4742541 A251836 AB893501 A1062568 AB34777 A665122 AA014938319 AA423082 AB483038 A2026751 A41134941 A1312660 A203095 A006798 H07076 AC25442 AS354542 A151169 AC332294 AW047480 A424735 AW047476 A17045 A4756160 N13567 AW051136 A119234 AC23410 AB919237 AS25846 AC258966 A421996 A137347 A139460 A1426583 H7944 A426991 W007157 A4474751 A348708 BE59668 AB463389 AW951951 T28593 AA458803 A307107 AA281033 AA029938 AW040805 F06483 BE41406 W03282 AA548291 N01028 AA701106 A516401 A575546 A1993030 AA481792 A1176983 AW845551 AA639142 A2171794 A1043733 A434591 AS927196 A274416 AA564231	
15				
25				
35				
45				
55				
65				
75				
85				
95				
105				
115				
125				
135				
145				
155				
165				
175				
185				
195				
205				
215				
225				
235				
245				
255				
265				
275				
285				
295				
305				
315				
325				
335				
345				
355				
365				
375				
385				
395				
405				
415				
425				
435				
445				
455				
465				
475				
485				
495				
505				
515				
525				
535				
545				
555				
565				
575				
585				
595				
605				
615				
625				
635				
645				
655				
665				
675				
685				
695				
705				
715				
725				
735				
745				
755				
765				
775				
785				
795				
805				
815				
825				
835				
845				
855				
865				
875				
885				
895				
905				
915				
925				
935				
945				
955				
965				
975				
985				
995				



			AN075390 A085809 A311089 W02203 A241526 T52777 AA092000 T52778 AA30823 A2731417 K21540 A000911 A1146570 A083741 A869897 A0505429 B0369067 D48151 A4861638 R36011 N3429 T03027
			AA163514 NM_004209 A085809 T08741 T04072 A08523 R1897 B3175369 H20003 K04661 A00157065 B08728 A0021563 J030873
5	112941	4686_1	AA12001 A104066 A4776642 A083308 T16322 H20614 A058007 A04378765 B0314614
	105726	5801_1	NM_012686 A022663 T1251 T07734 A43680 H06545 A101118 B02711 A553305 A0412936 T0677 T7368 D33114 A070651 A0020150 B020701 A1161154 A022308 A445133 A094746 A191059 A094766 N05770 A000315 H04824 W05638 T04035 W36601 A00157065 B0369067 T08741 A08523 A4776642 A083308 T16322 H20614 A058007 A04378765 B0314614
10			DE03348 A0601266 B0603943 B0604227 B063949 B063947 A06360 A06360 B06360 B06360 A061282 A345147 A01281 B08
			B063945 A053226 B064601 A05323 A059163 A033063 B063950 A053060 B06360 N03604 A048025 A0776302
			A0776302 A0003151 B063946 A0776302 A011937 A0421049 A0179101 T01305 A05506 A000002 A03002 A056815
			A087208 A345152 A262172 A0747221 A0887954 Z29406 B063941 A0720657 A076647 A218163 A094699 A072064 A1 A0173974
			A142666 H03560 A0654521 A132070 A077332 A2426309 A0303757 A024020 T03600 B07809 B063951 T07671 R00307 T05002
			A0776302 A0003151 B063946 A0776302 A011937 A0421049 A0179101 T01305 A05506 A000002 A03002 A056815
105731	18373_1		T03600 T45962 A0003151 B063946 T03244 B0175208
			A054604 B017604 A0177881 A063538 A02357309 A013061 N50490 F07014 A4464969 A342929 A000304 A035640 A002024
			T07995 A0160157 A0849973 A052063 A0060903 A336376 A557462 A070795 A000000 A0133734 A0379310 A016108
			A0241303 A0356417 A0455270 A035593 A0208035 A0235760 A069864 A033067 A0292711 H13206 A0470178 T07074 A0475790
			N505305 A053226 A066674 R77646 D21194 A07770026 A0208078 A038439 A038238 A058792 N00415 A352442 A032902 A032903
105930	7471_1		A053226 A0003151 B063946 A0776302 A011937 A0421049 A0179101 T01305 A05506 A000002 A03002 A056815
			A045812 A011055 A0702913 A017464 A0450954 A057746 A0621121 H06372 A0038346 A062079 A543070 A068213 A0737374
20			A024216 A085668 A0134915 A067321 F06364 A0391510 A220820 A066261 A081213 A00106591 A5329790 N072502 A0139444
			A083546 A081281 A1583077 A021047 A0574006 A047807 A088996 A072882 NM_005672 A045436
129265	2730_11		A053092 R04352 R22953 W51995 NM_004417 X08277 A0013022 R05900 A013774 A090147 A019479 A0459716 R22310 A0524040
			W15917 R02658 A1122673 A039254 A0450495 A038260 A0454911 W04005 H10201 A0406108 W05651 A000000 A03749 A07595
			A052800 A04149M A0303637 A0183973 A008149 A000726 A470067 A471284 A023492 A020605 A0315019 A088365 A032134
			A305942 A072472 A0635363 A521431 A004940 A4760212 A0389747 A0208830 A072280 A006713 A033238 A0418500 A0435473
			A046966 Z22641 A053719 A020361 A07421 A048784 A000473 A0432991 A045694 A013068 S3 A033632 A0317705
			A043633 A072274 A067096 A127815 A042473 A036811 A083063 A001947 A0039370 A003546 A003546 A0170689 A06828
			A071510 A0327367 A0182408 A144465 A068506 A0214509 A008065 A068279 A0456351 R06153 A068602 A041299 A03100
			A071816 A024381 A0835016 A0457094 A003369 A089177 A1212615 A0001870 A080287 A025372 A074510 R04308 A0471928
			A0010446 A020236 A0478176 A074036 A0665560 A472161 R25023 R069843 A049694 A14350 A214501 A007023 A191094
			A052158 A047126 A000000 A072295 A000000 A000000 A000000 A000000 A000000 A000000 A000000 A000000 A000000
			H07112 A19923 A06874 A03175 A000000 R22200 H1502 R7359 A004454 A067643 A0712547 A0891246 A052977 A076311
			A000000 A077220 A0991570 A071254 A078205 A031071 A214083 A063532 R06348 R32615 H08494 H07747 A052075 A017684
35			A094745 A053236 A061673 A0520640 A0791046 A0620662 A171332 A0565322 A087663 A003249 A0472827 R1 A061351 A041663
			A004774 A063086 A0475740 B06229 A19995 A000633 A0537061 A0682278 W0424 A004049 A0090707 A004763 B053760
			A054529 A053650 A0716181 A0041817 A021785 A007662 B023015 B04547212 A00964 A000000 A000000 A000000
			A0436301 A0422196 A019323 A00025158 A047340 A0422214 A01373028 B029550 A007383 B021930 A002046 A00414315 A0015004
			W44534 A047341 A151365 A0675231 A1436034 A039217 A076900 A0218794 A0130622 A058028 A0050728 A2103083 A033004
			A0022327 A058018 A053671 A008177 A081815 A000096 A000096 A000096 A000096 A000096 A000096 A000096 A000096
			A0130007 A0002971 A000000 A000000 A000000 A000000 A000000 A000000 A000000 A000000 A000000 A000000 A000000
			A0013006 W4453235 H0894 H05546 W083203 A0376151 A3323008 R4196 W12976 A075472 A008726 W22246 W1874
			R06537 A006723 W35757 R01631 A0754208 A0569883 A088308 A296041 A065272 A056076 A0448331 A0151060 A0149094 T121953
45			A0110578 R067131 A031268 A0673791 A0545819 A052834 A0485000 A041359 A068226 A0229714 A010043 A0534962 A058001
			R05871 H0421 A002323 A01150717 A0090232 R0459 A042111 A354466 A085249 A0132259 A051657 A037416 A055622
			A0571483 A013688 A03367610 A375927 A0564532 B0122573 T07941 R28390 A014909 A0073346 A005306 A0205165 A043403
			A051368 A148236 A045096 A0342413 A0118052 R06861 A0578126 W272437 A0947429 A072106 D2558 A070997 A073640
			A053584 A0418187 R20207 A0003003 A051476 A0432527 H05546 A063548 A0051042 A049850 A0100542 A0472058 A0476252
			A0040426 A0075947 A097949 A000594 R0368 H0574 A0150431 C05699 A0042914 R07358 A0405912 A005418 T03811 A0022644
			R07359 A038257 A02389 A073594 R0371 H0574 R05642 T06072 A00061 H0584 A041894 A047410 A034169 A001779
			R3249 A0473913 A0654386 R08738 A1244700 D6190 A072047 H04396 R0941 R05504 H06903 B024316 H0587 B0122905 A092624
			A033760 A040761 B008977 A00140876 H04938 R02399 R06087 T1257 F01408 A04484 H05400 A04022 W0037
128610	5906_1		N43731 NM_010127 L36504 Y10183 A022217 A073075 R03874 A062218 A000000 A044626 A002550 R1356 A0214664 A000000
			R05871 H0421 A002323 A01150717 A0090232 R0459 A042111 A354466 A085249 A0132259 A051657 A037416 A055622
			A0571483 A013688 A03367610 A375927 A0564532 B0122573 T07941 R28390 A014909 A0073346 A005306 A0205165 A043403
			A051368 A148236 A045096 A0342413 A0118052 R06861 A0578126 W272437 A0947429 A072106 D2558 A070997 A073640
			A053584 A0418187 R20207 A0003003 A051476 A0432527 H05546 A063548 A0051042 A049850 A0100542 A0472058 A0476252
			A0040426 A0075947 A097949 A000594 R0368 H0574 A0150431 C05699 A0042914 R07358 A0405912 A005418 T03811 A0022644
			R07359 A038257 A02389 A073594 R0371 H0574 R05642 T06072 A00061 H0584 A041894 A047410 A034169 A001779
			R3249 A0473913 A0654386 R08738 A1244700 D6190 A072047 H04396 R0941 R05504 H06903 B024316 H0587 B0122905 A092624
			A033760 A040761 B008977 A00140876 H04938 R02399 R06087 T1257 F01408 A04484 H05400 A04022 W0037
			N43731 NM_010127 L36504 Y10183 A022217 A073075 R03874 A062218 A000000 A044626 A002550 R1356 A0214664 A000000
			R05871 H0421 A002323 A01150717 A0090232 R0459 A042111 A354466 A085249 A0132259 A051657 A037416 A055622
			A0571483 A013688 A03367610 A375927 A0564532 B0122573 T07941 R28390 A014909 A0073346 A005306 A0205165 A043403
			A051368 A148236 A045096 A0342413 A0118052 R06861 A0578126 W272437 A0947429 A072106 D2558 A070997 A073640
			A053584 A0418187 R20207 A0003003 A051476 A0432527 H05546 A063548 A0051042 A049850 A0100542 A0472058 A0476252
			A0040426 A0075947 A097949 A000594 R0368 H0574 A0150431 C05699 A0042914 R07358 A0405912 A005418 T03811 A0022644
			R07359 A038257 A02389 A073594 R0371 H0574 R05642 T06072 A00061 H0584 A041894 A047410 A034169 A001779
			R3249 A0473913 A0654386 R08738 A1244700 D6190 A072047 H04396 R0941 R05504 H06903 B024316 H0587 B0122905 A092624
			A033760 A040761 B008977 A00140876 H04938 R02399 R06087 T1257 F01408 A04484 H05400 A04022 W0037
			N43731 NM_010127 L36504 Y10183 A022217 A073075 R03874 A062218 A000000 A044626 A002550 R1356 A0214664 A000000
			R05871 H0421 A002323 A01150717 A0090232 R0459 A042111 A354466 A085249 A0132259 A051657 A037416 A055622
			A0571483 A013688 A03367610 A375927 A0564532 B0122573 T07941 R28390 A014909 A0073346 A005306 A0205165 A043403
			A051368 A148236 A045096 A0342413 A0118052 R06861 A0578126 W272437 A0947429 A072106 D2558 A070997 A073640
			A053584 A0418187 R20207 A0003003 A051476 A0432527 H05546 A063548 A0051042 A049850 A0100542 A0472058 A0476252
			A0040426 A0075947 A097949 A000594 R0368 H0574 A0150431 C05699 A0042914 R07358 A0405912 A005418 T03811 A0022644
			R07359 A038257 A02389 A073594 R0371 H0574 R05642 T06072 A00061 H0584 A041894 A047410 A034169 A001779
			R3249 A0473913 A0654386 R08738 A1244700 D6190 A072047 H04396 R0941 R05504 H06903 B024316 H0587 B0122905 A092624
			A033760 A040761 B008977 A00140876 H04938 R02399 R06087 T1257 F01408 A04484 H05400 A04022 W0037
			N43731 NM_010127 L36504 Y10183 A022217 A073075 R03874 A062218 A000000 A044626 A002550 R1356 A0214664 A000000
			R05871 H0421 A002323 A01150717 A0090232 R0459 A042111 A354466 A085249 A0132259 A051657 A037416 A055622
			A0571483 A013688 A03367610 A375927 A0564532 B0122573 T07941 R28390 A014909 A0073346 A005306 A0205165 A043403
			A051368 A148236 A045096 A0342413 A0118052 R06861 A0578126 W272437 A0947429 A072106 D2558 A070997 A073640
			A053584 A0418187 R20207 A0003003 A051476 A0432527 H05546 A063548 A0051042 A049850 A0100542 A0472058 A0476252
			A0040426 A0075947 A097949 A000594 R0368 H0574 A0150431 C05699 A0042914 R07358 A0405912 A005418 T03811 A0022644
			R07359 A038257 A02389 A073594 R0371 H0574 R05642 T06072 A00061 H0584 A041894 A047410 A034169 A001779
			R3249 A0473913 A0654386 R08738 A1244700 D6190 A072047 H04396 R0941 R05504 H06903 B024316 H0587 B0122905 A092624
			A033760 A040761 B008977 A00140876 H04938 R02399 R06087 T1257 F01408 A04484 H05400 A04022 W0037
			N43731 NM_010127 L36504 Y10183 A022217 A073075 R03874 A062218 A000000 A044626 A002550 R1356 A0214664 A000000
			R05871 H0421 A002323 A01150717 A0090232 R0459 A042111 A354466 A085249 A0132259 A051657 A037416 A055622
			A0571483 A013688 A03367610 A375927 A0564532 B0122573 T07941 R28390 A014909 A0073346 A005306 A0205165 A043403
			A051368 A148236 A045096 A0342413 A0118052 R06861 A0578126 W272437 A0947429 A072106 D2558 A070997 A073640
			A053584 A0418187 R20207 A0003003 A051476 A0432527 H05546 A063548 A0051042 A049850 A0100542 A0472058 A0476252
			A0040426 A0075947 A097949 A000594 R0368 H0574 A0150431 C05699 A0042914 R07358 A0405912 A005418 T03811 A0022644
			R07359 A038257 A02389 A073594 R0371 H0574 R05642 T06072 A00061 H0584 A041894 A047410 A034169 A001779
			R3249 A0473913 A0654386 R08738 A1244700 D6190 A072047 H04396 R0941 R05504 H06903 B024316 H0587 B0122905 A092624
			A033760 A040761 B008977 A00140876 H04938 R02399 R06087 T1257 F01408 A04484 H05400 A04022 W0037
			N43731 NM_010127 L36504 Y10183 A022217 A073075 R03874 A062218 A000000 A044626 A002550 R1356 A0214664 A000000
			R05871 H0421 A002323 A01150717 A0090232 R0459 A042111 A354466 A085249 A0132259 A051657 A037416 A055622
			A0571483 A013688 A03367610 A375927 A0564532 B0122573 T07941 R28390 A014909 A0073346 A005306 A0205165 A043403
			A051368 A148236 A045096 A0342413 A0118052 R06861 A0578126 W272437 A0947429 A072106 D2558 A070997 A073640
			A053584 A0418187 R20207 A0003003 A051476 A0432527 H05546 A063548 A0051042 A049850 A0100542 A0472058 A0476252
			A0040426 A0075947 A097949 A000594 R0368 H0574 A0150431 C05699 A0042914 R07358 A0405912 A005418 T03811 A0022644
			R07359 A038257 A02389 A073594 R0371 H0574 R05642 T06072 A00061 H0584 A041894 A047410 A034169 A001779
			R3249 A0473913 A0654386 R08738 A1244700 D6190 A072047 H04396 R0941 R05504 H06903 B024316 H0587 B0122905 A092624
			A033760 A040761 B008977 A00140876 H04938 R02399 R06087 T1257 F01408 A04484 H05400 A04022 W0037
			N43731 NM_010127 L36504 Y10183 A022217 A073075 R03874 A062218 A000000 A044626 A002550 R1356 A0214664 A000000
			R05871 H0421 A002323 A01150717 A0090232 R0459 A042111 A354466 A085249 A0132259 A051657 A037416 A055622

PCT/US02/17594

1















WO 02/098358

PCT/US02/17594

[illegible]

100452	31861_1	D67742 AL041819 AB00015 AB0592074 AN740413 L36488 AB020000 Z29572 AW050061 A208678 BD011008 AL040523 AA036396 BF17291 BE025789 AA231074 W03622 AA233892 AA236265 AA368813 BE083675 BE098991 BE020532 BE082002 AA381354 AA231995 AA381648 AA411613 A135917 AA257960 AW048480 AW048281 AW001576 AB144678 BA063356 AA357375 AB078885 AA444636 AA369608 AB027279 AB454860 NS3228 AB179021 AW097077 T7841 AW020698 RT1379 AB085944 AA972682 AA0369873 AB02205 AB27641 E31156 T71944 T70949 AN07178 BE005352 AA025006 AW047428 AW062311 W13775 AW001225 AA411153 A240391 AA058575 AA025136 AA002517 AN751601 H15021 AW023254 AA053915 AN591728 AB19780 A077846 A030710 A0269162 H257533 H74270 A219154 AB072714 AB178958 AB063084 AB300750 AW058690 H67534 AB02375 72830 J32629 Z41948 A026434 N98415 W36480 R00445 AA319032 AW562141 H69052 AB000042 NS2690 A0040740 AA432213 BE467872 A210115 A412552 A001367 BE47054 AA249202 BF216989 H15445 AA030450 AB097639 B0716510 A008771 H08053 H19653 H23085 AA757277 H40250 A112093 A0040690 A0008042 R00449 A205452 A028605 H2626 A1015772 AW05841 AB073626 T02215 W37070 A034477 AB068645 H40487 Z38236 A01224 D3176 BF01641 DC0371 AW058422 AW089416 A0575383 F062460 AA043947 AW02820 A075247 AF95287 AB038676 AW057275 A0033696 AA430241 A2419740 A0134552 A000764 AA330325 A075242 AA33699 NS0384 AA477963 BE054942 A005511 H40488 AB075000 A0171797 AF098299 AF098289 AB069781 A1770170 R03502 A081272 A005738 BA28572 A075119 H19357 H08453 H19782 A023678 AW051984 A070444 AA051555 A018611 A107318 AW174567 AB033549 AB081273 NS0408 AW078138 H04066 AA070677 H77519 AB074047 AW057580 AW073073 A0014035 A0178876 A037954 AA170520 A1107863 AW027292 NS3693 A1192508 BE119405 A1140076 AB098555 A0861023 N03319 J15244 A1147652 AA812948 A1442325 A1367868 AB093607 R03114 AA430287 AA287965 AA1770126 A020002 A0043585 AA172482 A4059412 A04120 A0015240 NS5467 R03011 H77520 A027854 AA043002 AA055120 H08693 A110594 H67910 R03255 H61512 AW023666 H04552 AA0409296 H60307 H60733 F05322 AB067719 H02547 A074762 A080902 A033075 A0267547 A0405532 A0023620 AA04064 H02417 AB057325 C1248 A037066 BE32465 H25411 BE261824 A0072365 H06115 H02577 A0352192 H00872 H67960 H66303 H67633 H52308 AW176097 A1719814 AW196237 AB11894 AA044227 A232363 A0020478 AA383695 AA040193 A0015339 AA359435 BE113893 BE230644 BE281929 A006140 A0061435 A331949 H010936 A0042232 A1002994 A0036322 A0047470 AA459475 AW07400 W07878 AW073653 BE13241 AA465099 A1046680 A1343696 AW003264 BE301261 A1707897 A0198487 AW576703 W06453 A0004688 AW076701 AA223724 AA366963 A0215565 AA44565 AW193840 A0196227 AW07789 BE014036 A0178876 H183406 A0359331 AA455044 A128996 AA058705 A113839 A0176292 A05152191 A0242425 AA452924 A0602279 A0614036 A1459782 AA42824 A1435481 AW075178 AB094425 AW075178 BE27460 A026793 AW054089 H63744 A0236910 A4729135 AA766112 AA155335 AA353092 A0751398 A1767126 H01_002947 L07483 AW17080 A056682 A0404079 A0232077 A141138 A174142 AA41254 BE055628 A0334741 AW110993 AA345343 A0671710 F05576 A065229 A0456226 AA374637 A0646302 A3374248 A247681 A190616 AW0718005 A350208 A1351115 A082467 H09259 A250796 A001330 A243725 H4361 A018597 A024368 A1241212 A017864 A0187429 A19074021 W03604 A0907487 T27775 F36677 AA743038 AW049944 AA049985 A0472785 A080111 A247562 A0103673 A0130747 A0751917 A1430702 A07517540 A0710533 A0264398 A4227780 A1735761 A035148 A0705038 A087055 N03011 R02112 A0151689 AB055004 AW381977 A015678 A0017653 W06843 A0225138 T06667 F0261 AW138353 A40105242 AA040296 A013431 AB08980 A0460002 AW40424 R30965 M09241 A172220 AW161570 BE311351 M35410 BE325458 S37730 AA383048 A040345 BE295221 A47180343 A1722261 BE314091 BE32585 A0308676 AA327814 BE26917 BE332555 AW035824 A021670 A00180345 T61701 A001435 A015794 A0189226 BE253473 A021677 BE26162 A049155 AA45833 AB07353 A0440417 A069011 A4373839 W02328 AW300897 A015789 A01727441 A0219659 AA419054 A1143021 A052277 A0780788 AW161913 A0320681 A062804 AW374612 AW161619 A080911 A0354468 AW157805 AW138353 A0183887 A0166904 A0383011 N3178790 A0663332 AW0593632 A042811 A118554 A134101 A0180197 A014181 A081026 A1143921761 AA435454 A018085 A087263 A005351 A412300 A039444 A065301 A0170870 A0284142 A1783893 A071958 A0302749 AB336910 A001554 A000300 A030680 A0430001 H75047 A003571 BE00161 A0123788 AW0727441 A026476 A1725280 A1752221 A0172776 A102334 A05113739 AW059620 A0035641 A0651216 A015236 A00082 A069242 A047639 AW112524 A0503462 AW09096 A004461 A0401907 AW059316 A185447 A175115 A1500245 A064378 A336842 A1872625 A01192270 A066555 A0331338 A0445494 A230921 A024935 AB04402 A0404300 AA477419 A417952 A136212 A1028374 A1567133 A017582 A020475 AW513319 A0513432 A051185 A052123 A077119 AW052056 A016670 A0151048 A4077269 A05345 A01152162 A355468 BE304794 A355695 A171487 A0133290 A071987 A077441 A240502 A0735946 A0448641 T26252 A669991 A0029590 A16302 AA592329 A2382780 AW062374 AN051577 A0402821 A037725 A176575 A174571 A0111904 A015698 AW514321 AA350379 A053278 A082486 A0791194 AB21390 A0504794 AW099151 F37127 A054206 F27974 A381807 A47144755 H01_013332 A0320007 BE294300 BE311752 A0301320 A0501220 BE30104 A0100427 A309121 A1343669 A004057 A005452 A1310162 A0744970 A769640 AW074287 AA461187 AW116166 A1277620 A0417821 A003341 A067014 A174879 A760482 AW51480 A054915 A063906 AW105694 AW010638 A148588 AW03071 BE454900 A0082974 A0628078 A0236191 A01241771 A038420 A332595 A0766796 AA054930 A0514387 A37851 A040468 NE3714 AW196666 A0354649 AW063546 A004300 A0979736 A035470 A056846 A002158 A0170824 A1743202 A193509 AB31483 BE464933 A0330387 A1939106 A465473 T2489 A016585 A014688 A44952 A092579 A019781 A134620 A045915 BE350612 A474514 A476709 A4775633 A0614887 A294616 A451953 AW176994 A17031659 A03545 A1865521 A1922877 A0173650 A040348 BE1950 A0331456 A025251 A372701 AW086592 A067283 A01693305 A332405 A0736861 N05719 A331393 A0468183 BE06946 A0644370 A053816 AW472739 AW43820 A0439625 AW0520271 A062973 A0551172 A464726 A463944 A225454 AW066734 A001468 A4190315 A347480 AW061179 A3037782 A3145296 A0437194 A0038690 AW036819 A01398192 A07722 A07124 N05976 A215522 A0263699 N07535 R12261 NS0794 AB060045 A347193 R1612 A1711 B0006 N05995 N06706 A07617 A01261 A06514 D0205 A175185 A030716 F04026 A112927 A1133290 A063807 A0440417 A077441 A240502 A0735946 A0448641 T26252 A351508 A072769 A091906 A11216 A0001472 BE058761 A058104 A006820 A070025 A0645493 A0090911 A0070815 A000361 A025930 AW29690 A000490 A004642 A0296253 A0049701 A481387 A425101 A422024 A011871 BE069191 A00611 A031354 A15443 F06442 A00801 H03310 A001320 A0501220 BE30104 A0100427 A309121 A1343669 A004057 A0053737 AW12263 A00111 A063871 N06866 A038726 A03463 A060477 A031510 A0073920 A0179921 BE330961 A156133 A01367 A004954 A139425 A0810261 BE20003 A01367616 A085210 A10387612 A03367610 A0509751 A0357527 A0378768 AW387632 AB14324 A0007818 A240600 AW097814 AW037624 BE03967 AW387586 A0375256 A037879 A007949 A0387601 A1021751 AW08070 A002046 A01367516 A0387516 A0387516 A0387516 A0387516 A0387516 A0387516 A0387516 A0387516 AW387631 A0387636 A0387633 A0387635 A0007739 A0007695 A0387677 A0387694 AW07817 A0387690 A007690 A0387651 AW387521 A0387602 A0387625 A0387666 A0387657 A0387650 A0387650 A0387650 A0387650 A0387650 A0387650 AW387596 A0387635 A0387667 W1167 A020007 A0387669 N09676 A0176802 A017683 A007069 A214682 A0402794 A0387674 H17425 A003690 A003690 A023281 A0422284 A00876 A0387627 A0387544 A0387591 A0387702 A0387610 A0581875 A0387687 A0387696 A0387655 A11115 A0387636 A026819 W02874 BE14948 A0029824 A4115298 BE048130 T09421 A0358287 N053397 A352571 A174284 A016077 BE039130 A425353 A4464842 A0189900 A031487 A078218 A120709 A0387420 A0038746 A127391 A444442 A0025486 A0857533 A008647 A031579 A025258 A007626 A1130618 A008761110 NS4862 A0086698 A0086698 A0086698 A0086698 A0086698 A0086698 A0086698 A0086698 A053183 A0430526 A007691 A007693 A00231 A293396 A421006 H47336 A110937 N06163 A0081085 A0064325 A042056 A0092723 A0272436 A0995657 F22574 H09012 A021968 A008361 A4544167 A037392 A0321703 A0081085 A0064325 A042056 A0092723 A0272436 A347546 A0337894 A299197 A025355 AW067641 AW026268 A051794 A018119 A0191521 A12421 A094720 A034047
--------	---------	--







WO 02/098358

PCT/US02/17594

5

10

15

20

25

30

35

40

45

50

55

60

65

70

75

80

H20505 AA635234 AA857696 AA026821 AA026740 AA036356 AA045893 AA100369 AA133047 R00112 AA101759 AA159963 AA026804  
AA191988 AA122445 AA112416 AA155634 AA066612 T49519 AA657980 AA023082 AB926520 AA315544 AA155647 AA972452 AA811776  
AA009165 AA076555 AA531341 AA071963 AA070630 AA189148 R00874 AA076125 AA020434 AA056910 R02091 AA145650 AA157271  
AA113161 AA061386 AA185270 AA113477 AA033861 AA073513 T53238 AA035655 AA153366 AA035655 AA157857  
AA041842 W06455 AA007325 AA158304 H07723 AA137353 AA126335 AA100381 AA121070 AA455226 H12796 AA077667 AA054059  
AA563680 AA079134 AA174181 AA840642 AN272916 AA161228 AA011735 H01810 AA020286 AA158779 R01923 AA034689 H27392 R02239  
T501302 AD079799 AA069298 AA006120 AA534348 H01719 AA847550 H00189 H026470 T29401 AA124943 AA077856 AA123123 R01594  
AA056462 AA0264947 L20014 AA155267 AB032407 AB04045 AA54046 AA20108 AB073088 AA032338 AA000202 AA053715 AA111570  
R01512 AA062321 C00953 J00001 AA051420 D06558 AA367222 AB01152 AA005120 AA056467 AA176151 AA1158029 AA115978 AA1122412  
AA0369925 AA075964 AA143006 AA101052 AA122141 AA150526 AA143409 AA158166 AA100035 AA007930 AA007876 AA367322 AA134136  
AA101364 AA126355 AB0328 1W231 1A9518 R31821 AA071965 AA026434 AA05701 AA063100 A12201 C06137 AA367372 AA552340 R32468  
R00875 H20901 R70146 H028454 AA133466 AN270871 AA026236 AA026795 AA059992 AA0081676 AA012314 AA051745 H00575 H140272 AA026639  
AB07195 AB033003 AA041030 AA027590 AA047659 AA070563 AA027046 AA057038 R01715 AA145265 R02066 AA133037 AA151730  
AA007347 AA422246 AA009504 AB06815 AB06817 H00919 AA158004 T52627 AA075536 R00001 AA069617 AB01163 AA159069  
122860 genbank\_AA64414 AA046414  
109466 genbank\_AA079409  
108506 genbank\_AA063076  
entrez\_M11321 M11321  
109679 genbank\_AA115963 AA115963  
124367 genbank\_N22401 N22401  
101544 entrez\_A51169 M3169  
124777 genbank\_R41933 R41933  
117789 genbank\_N48294 N48294  
119071 genbank\_R31160 R31160  
133512 9798\_5 119861 L16685 AA0070431 AA113637 AB032547 AB16703 AB073940 AB020667 AN200642 AB0073179 AA113124 AA143432 AN201718  
AA104933 AA050446  
AA134233  
105057 genbank\_AA134233  
103804 AA129196\_at AA129196  
135905 19622\_1  
AD028774 T23531 AA311697 AA198930 AA302066 AA322667 AA004830 AA027601 AA136577 AA333383 AA323996 AA333306 AA061713  
T06471 H033338 AA11162 F02011 AA033338 AB00889 AA026555 T07449 AA13621 R01641 H03712  
AD02177 W79748 R07255 R07252 W06358 AA303024 AA326988 Z20422 AA100131 AA330349 AA102098 W07556 AA11689 AB029519 W04646  
R06040 AA027771 R45272 W73894 F11842 AA122604 AA143062 W06745 AA127271 AA150826 T00201 Z45719 AA026367 T35161 H00507  
AF038197 F10516 R16648 W05356 AA150600 AA11688 R44590 W72861 AA112723 T30476 W06096 R401465 AA119593 T05874 T21466  
AA019154 Z40126 W06356 AA0137027 W01401 AA129676 R64088 AA006067 AB073946 R27258 W05339 N42343 Z42761 H52474 H05176  
AA078567 AA078562 AA078568 AA078566 H0275651 AA179538 AA075965 AA075961 AA075960 AA075959 AA075958  
H00293 N1738 H03037 R70566 R03825 C01169 H51765 R02732 AB021611 N64724 W09455 AA119831 AA328272 AA827256 W06811 H037519  
AD036882 AA077075 AA126955 AA125094 AA367630 AA367611 N64724 W09455 AA119831 AA328272 AA827256 W06811 H037519  
AA074083 AB049452 D61614 H02762 H00832 H00831 W06200 W07326 H08874 AB026598 D61742 AB02100 W069145 R06841 R069791 N34686  
H02523 H05058 W09962 AA011307 H01435 AA046572 AB04410 R03465 H09406 AA07041 AA183561 N03365 N02016 AB08115 AA130298  
AD030422 H07584 AA067587 T34001 H08073 R75987 AA047060 AA029932 AA025123 AA191684 R17603 AA0510720 AD02832 AA056717507  
AD004242 AB0530 R02477 AA129838 R41677 AA100115 AA021449 R31444 H01366 F11606 AD244779 H03041 AA666067 AA003460  
AA079680 F00874 R36520 T34900 AA831189 AA033886 R05643 R76118 T31573 T35052 R21448 T30304 AA032236 AA1163443 AA0470784  
Z41376 H02603 AA302269 AA074002 AA020922 AA222222 H08090 R02090 AA02061 W061141 AL036961 W07653 R00730 AA304264  
W095101 R000761 AA416931 AA416942 AA057879 T029865 A113437 AB003495 T04060 U21458 R05332 AA002713 R71765 AA307660  
R14227 R14543 W09251 W72546 R70477 R24022 H07946 W0116 W55469 R03564 R70593 R03656 H03653 R06044 T04346 R78115  
R78127 T33467 H04100 R73790 T34997 T04579 AA0206275 F11716 H08311 N02211 F13516 H05691 AA307966 R07496 R07496 R02212  
AA033461 R10023 R09712 AA327621  
116921 NOT\_FOUND\_entrez\_W03038 W03038  
116546 NOT\_FOUND\_entrez\_W38169 W38169  
118559 NOT\_FOUND\_entrez\_W38197 W38197  
128046 877065\_1 AA872266 A025762  
135424 US0111\_at US0111  
126480 genbank\_T12608716206  
114767 AB09965 AA148885 AB05593 AA0701342 W74071 AA211366 AA050010 AA041939 AA0470717 AA508102 T099175 AA379782 AA373531  
105047 igr\_H72219 M57417  
323035 33334\_1 AL137513 BE072892 AD127078 AA0106207 AN294979  
351408 W07930\_1 ANK61530 R090402 A234443 AA0125077 AB090404  
321412 AB024692\_1  
321415 62865\_1  
AB073483 AA065710 AD104051 AB080387 U25919 BE093109 AA050305 BE141926 BE141913 AA054334 AA0854342 AA0415916  
AB071625 AA145481 A1346385 AA037413 AA164664 AB028364 AB046527 BE109214 AB094111 AB091352 C17504 C17476 C17693 C18304  
AA071625 AB07162 C17732 D05758 H01762 AD120938 AD020590 H27262 A44792 T2 AD027868 A291466 AB155661 AB093432 AB057698  
AA424176 AB052667 AA022666 AA119248 AA119249 AB026338 AB069605 AB119040 AB119121 BE011941 AA432916 AA161916 AB09761  
AH24304 AA05610 AD140098 AA020230 AA025254 BE044033 AA026237 AA006510 AA043216 AA026595 AA138646 AA03772 AA070657  
AA08439 AB006193 AA001803 AA479834 B05019587 AA129574 R08114 AA049484 AA024628 BE032120 AA057253 BE021874 AB039186  
AA015724 AA043217 AA077200 AA079841 AB017127 AA0778725 AA020832 AA043093 AA0129597 W08161 AB07219 AA046570 AA027628  
AA123200 AA147225 AA030261 AB002406 AB075678 T00407 C06123 AA15794 AB007036 W00075 A103915 AA013326 AA057352 AA1182640  
AB090927 AA0275048 AA0010340 A0285955 ANW1421 R71990 AA005159 W45410 A033371 AA0030456 AA062517 T55841 AB023466  
AB028465 AA062397 AB091997 AA136568 A0251817 BE04134 AA033914 AA071762 AA27439 R79533 AA411100 AA191349 AA007696  
AA190956 AA175735 AA077283 AA010631 H008043 AB070916 H049696 AA001065 AB005693 AA003842 AD245632 BE023694 AA009386 BE150360  
AA039380 AB021320 AB138844 AA019472 AA146284 AB003683 C03636 AA467093 AB072171 AB185253 AA043465 AA078602 AB19396  
AB021369 AA022660 AA033616 AB072539 AB026754 BE072626 AB093000 Z08711 AB056593 AA053209 AA033901 AA681093 AB050336 AA003632  
AA003636 AB098240 AA024296 AA119173 A174517 R1 AA067300 AB024662 AD24645 AA056439 F00579 AB116453 AA025575 AB12527  
AA031558 AA191414 AA061745 AD2558 AA06740 D07645 AB098064 C05782 AA072206 AA032003 R21752 BE157010 AD129640 AA027623  
AN054233 AA026982 AB06912 AA017954 A0345441 K75804 A034945 AA058310 AA090297 AA034660 A346877 AB094115 D11521  
AB077596 AB087665 AT12433 C08111 AA043607 AA069646 A274849  
AA062331 A1341081 AB048026 AD081645 A1368235 AA002023 C036901 A0374947 AA080781  
A107493 AA028275 A0350197 AF107492 AA007797 AB016216 AF142678 19232 Z20069 AA001646 AA029763 H74105 AA046626 AB33432  
A218239 AA170303 AA020615 AB061844 A0347061 AA059359 AA41684 A004197 AA042253 AA003709 AA042253 AA003709 AA042253  
AA057384 A03709  
AK09177 H70146 H70145 AA0392474 BE007303  
A025575 AA063432 AA038876 AA0578281 AA064986  
AA020691 AA063900







WO 02/098358

PCT/US02/17594

303274	61_1	A001458 AA190315 AA374900 AW961179 AA307782 AA315295 AA3407194 AW953073 AW368192 AA280722 AA251247 H85675 A012522 AA213838 AA300512 AA320542 AA330052 AA360356 AA362421 BE118820 BE207088 AW247908 AW328143 AA47165 AA300769 AA06261 AA122827 AA313250 AA02488 AA090372 AA888645 AW009617 AA24385 AA190314 AA173569 AA971106 AA051088 AA072789 AA05955 AA11215 AA001472 BE56673 AA559104 AF702623 TB027 H08911 F35413 H1048 TB0011 C54542 TB0144 F03822 F03030 ZV5148 R18285 AA354331 DE545752 AW229552 AW009717 AB35471 H08331 BE217766 AB372383 AW131702 AA21225 R49210 R03766 AA077779 R0875 AA08410 AA183124 AA04374 R14143 BE19C300 R56996 AA09152 AA486878 AA04573 AA356367 BE06584 AA079122 R07789 R5519 AW AW196387 AA10398 AA563719 AA17852 AA1280050 AA105042 AA125814 AA130052 AA063356 AA562421 BE118820 BE207088 AW247908 AW328143 AA093777 AA091450 AA052628 AA075223 AA070656 AA044241 R4834 AA638374 BE118310 AA056180 AA046427 AA030179 AA01411 AW195711 T0687 AA267540 AA34538 AA0967773 AA32289 AA51003 AA09850 AA039332 AA304483 AA05269 AA178205 AA435231 AA027675 AA152645 AA0751 AA1026395 AA168615 AA528979 AW153496 AA214278 AA02745 AA05679 BE567402 AA059355 AA245642 AA0744667 AA285441 AA07009 AA219314 AA211449 AA202567 AA043129 AA265302 AA21452 AA7515 R10E2873 NM_014785 TB47 BE253434 AA000883 AA407881 AA105015 N6150 AW058365 AA44421 AA272402 BE019777 AW199872 AA202685 AA102826 AA076801 AA105434 AA055032 AL133780 AA028914 BE548610 C31490 AA04391 AA056240 AA179059 AW117479 AA3411531 AA54611 AA274636 AA313732 AA255266 AA571392 AA052971 AA176369 AA053280 AA01125 AA1772523 AA361106 AA0683216 R45844 AA369352 AW236104 AA073669 F15747 AA02185 AA05810 AA19573 AA074612 AA143526 AA095236 AA2145439 AA091599 BE170801 BE13405 AW207562 AW92335 AA632616 BE040857 AA43198 AA18839 AA30777 BE274932 AA910729 AW051094 AA439156 AA335611 AW204095 AA195165 AW1977064 H94000 AA20991 AW567646 AA251451 Z45131 R20602 AA011796 AA234020 AA232982 H29165 T23514 AA55785 AA81545 AA01714 AW570357 AW073288 AA036336 AA359896 AA099790 AA73477 AA951615 T07547 AW304739 AA114041 BE176629 Z44590 T30422 T32690 AW95306 H10602 AA656662 AA014514 AA698482 Z45682 AA282123 H10149 AA050157 W05511 N7341 AF750919 AA01674 AA07700 AA08107 AA08602 R06252 R06349 F63159 R06301 AW170582 AA319967 BE122902 AA093462 T36150 T20223 T34001 H06172 F13007 BE515952 AA330336 T34747 R54849 T36254 AA128160 H04119 AA322989 AA342626 AA15125 AA13225 AA13217 TB0239 AF070648 H79087 AA748115 C02957 AA365970 H25456 H46655 H81233 R54555 AA003618 R46014 R43837 BE075503 BE151587 AA322855 BE021052 AA5373 W09034 AA155883 AA044383 AA096052 AA439387 N93600 R03256 R83439 W57400 AA461434 AA493873 V94189 AA054716 AA115260 AA556674 C03770 C02719 C02874 R05287 W03628 AA438399 AA429987 AA052695 AA373501 C03235 AA037146 H40914 C01252 AL13359 H02749 AA05510 AA151524 AA30467 AA148992 AA175400 AA330449 T81041 AA0477 AA35224 BE122903 R77190 R77045 AF074963 AA034379 R10132 R3576 AW196180 AA07741 AA089321 D61145 C04231 AT170909 AA304655 AA08307 AA007394 V44584 AA045458 AA60784 AL128025 AW005016 AW280208 AA034430 AA113602 AA008091 A333597 AA146439 N75560 AW578136 AA037268 AA77324 AA262326 AW574291 AT150990 AA41117 AA875326 AA503225 N09728 AA469355 AA373803 AA14953 AA00769 AA34390 AA073595 AA169947 AW0762 AA09277 AA033366 AA22229 AA08600 AA484102 AW095116 AA049359 AA018356 AA057560 H07122 M49761 T36254 R98990 H7252 N55688 AA657448 AA318382 AA43379 R30610 T07590 AA053731 T91713 AA095010 H9215 R76885 AA05319 AA61847 R07587 AA06879 R65688 H32062 AA080759 AW106862 R07401 AA487124 R56108 AA190163 AA104777 AA075533 AL243336 R96370 H61952 R64556 AA070904 H48499 R36343 AA37135 R081791 T57730 H22282 AA52224 AA043398 AA02848 R0827 R0827 R0827 AA57771 AA02653 T98237 BE452454 AA259331 AA133718 AA133718 AA447842 AA719834 T10086 AA302214 AA240527 D61798 AA242963 AA481755 AA004391 AA028905 AA142460 AA178065 N24039 R03163 AA052166 AA740015 AA0589311 T70254 R29483 AA679545 T96387 R8340 AA79020 T96698 C22616 AA044994 AA55758 AA172248 AA911537 H03416 T33767 AA013807 BE16545 AA615451 BE331199 BE461857 AA689127 T77611 T6303 R98872 H78291 T92749 R07586 R07540 T90281 N98435 AA030978 R39435 AA7910 AA052135 AA50276 T01066 AW11800 BE070771 N58545 C02523 C02563 H01168 AW08570 AA072655 AW082324 AA042650 AA110225 AA138225 AA62705 AA78119 N94916 H02931 AA221338 N04913 H24872 AA232915 AA522464 BE089972 A127509 R01404 A055314 AA279297 AA139622 AA074081 AA03034 N92762 AA355417 AA011285 R07413 R06778 AA9141 AA27432 N52608 BE004159 R191372 V94420 AA129453 BE166276 AT176990 AA048768 AW052146 AA572751 AA004720 AA246152 AA142874 BE000402 AA196248 AT1766610 V02394 AA232633 AA3175584 AA47817 AA183293 AA493414 A071114 BE16252 AA04357 AA04070 AA58583 AA061748 AA75178 AA04592 AA022066 AA032179 R1875 BE040242 DE001776 AA001633 A1220671 AA0812515 AA0812515 H07525 AA679008 AA050210 AA390657 AA005090 AA443524 AA052695 AA264499 AT176630 R20816 AA472739 Z40297 AA092827 AA026919 AW009737 BE044883 AA387303 AW275901 AW007987 AB043009 AA06862 AA304265 AA150227 AB72036 AA030999 AW11546 AA026420 AA005075 AT170306 AA504360 AA001961 AA132681 AA564285 AA061856 AW1951657 AA023635 AA57412 F36341 AA572391 AA394863 F22469 AA55164 AA671806 AW1750 AA07175975 AA561635 T30290 Z43131 H15301 A1565071 AA521568 AA093566 AA84217 AA058208 AA056726 AW198341 AA34005 AA048626 AL10225 AA34005 AA059241 A659190 AA07829 AA716150 AA27322 AA27434 AA072247 AA57045 AA53342 AA559005 AA916550 AA167602 AA671915 AA973048 AW252680 AA381875 AA92365 AA51544 AA072255 AA10285 AA065734 AA356317 AL17438 AA00741 AA348105 AA381079 H01955 AA286473 BE382495 AA296110 AW673101 AA315135 AA311617 AA325750 NM_00270 AA345143 X00737 BE265250 BE265212 AA375804 AA046329 T86291 T49631 N13963 H02039 AA27989 AA027997 T83203 N98627 BE379254 H37842 AA4527308 AW499647 AA148134 H83220 AA031946 W06334 W00892 R89169 R90427 H47282 H41854 AA343094 AA045099 R30316 AT102222 AA0274905 AA095968 AW078734 BE270064 N731 AW170022 AA06506 AA194872 AA65244 AA08493 H7211 AA1283427 AA0770 R05983 R2463 H13849 AA179319 H7006 AA019082 AA017006 AA42416 AA02183 AA02183 R05939 AA017006 AA017006 AA026889 AA053321 R66631 AA273191 AA572405 AA073735 AA026545 AA571872 AA073927 AA119078 AA21490 AA055215 AA37934 AA136181 AA140048 AA430082 AA04151 AA024764 AA549099 AA548275 AA31420 AA57495 AA14945 AA472559 AA8724158 AA043938 AA281952 AA279925 AA242703 W09290 AA013565 AA051176 AA14822 AA06057 AA043042 T47954 AA093504 R5017 H03996 W0262 AA077404 R08084 AA076114 AA020171 H02025 AA070746 AA075746 H1716 R02822 H05747 AA471332 R06832 AA18832 N6454 H17557 H8194 T95122 W08453 AA431099 AA431099 AA51574 N5361 AA057367 AA574743 AA4444 AA21475 H72533 AA799145 R93335 H0413 AA156105 AA209937 H04029 H89726 R58819 AA87331 AA1074096 AT35731 H02827 C02447 AA478005 T27051 AA99770 AA025156 N55225 AA063677 AA05965 H03719 AA428919 R22404 H13843 AA07376 R0088 AA03364 AA044142 AA541792 AA00141 AA128348 AA045090 N90434 T51267 BE1241751 AA0890720 BE38521 BE26557 W0006 AA071506 R9594 AA05484 AA087417 AA095190 AA070883 W00508 AA0116352 AA119216 AA037534 N55889 AA179470 W07304 AA1467 H83389 F12072 AW387905 AA054341 AA152326 AA071404 N31438 AA582293 T79422 AA019905 AA059312 AA218967 AA951481 T91861 AA053690 R06589 T7958 T84729 AA1967463 AA373019 AA039258 AA244432 BE06529
--------	------	--

WO 02/098358

PCT/US02/17594

5	311422	270835_1	FOC677 AW653696 AW061342 AA383426 AW966402 BE180650 AW041639 AW242125 AA060514 BE3011605 A1204986 BE219291 AW60015 A1274567 AW72210 Z2B533
	311465	857586_1	AW759590 AW759595 AW020610
	303654	64558_1	BE248743 AA436942 AW024744 AW242177 AW947546 AW369185 RW363 RW7462 AW654529 T57442 AW396660 RW0773 RW8743 AW789689 AW363655 AW719369 AW789001 AW189663 AW86207 AW471273 RW2345 AW315014 AW89181 AW896257 AW645044 RW05914 TW141 AW855793 RW0074 AW703263 A1271945 AW224459 AA505028 AB21061 AW51946 AW919161 AW706992 A2207290 AW568191 AW563075 AA333960 AA333562 AA24272 AW1567262 AW600040 AW563687 AW62452 AW196164 AW8744
10	310584	740029_1	AW041454 AW779953 AA434594 AW275353
	312105	413562_1	T81618 AW383709 AA703541 AW370105
	312108	202163_1	T82331 AW573065 HW3792 AA153717 AW51465 AA290850 AW270500 AW212726 AW293000 A1296364
15	312197	280163_1	T96203 AA405343 T96121
	311557	15183_2	AW52254 AW190862 AW102765 AW836668
	311598	903700_1	AWC25396 AW129233 BE360312 AW65034
20	310865	719002_1	AW177323 AW502030 AW949553 AW110381 AW859675 AW71924
	319408	301200_1	AA443090 RW5172
	303762	9344_1	AF034799 NM_003625 AW002917 AA325780 AW6544 HW8354
25	3115814	1702210_1	W07351 RW36746 Z42619 RW19469 RW800
	320099	21462_1	AW111007 AA336114 AF001555 A1223728 T31598 AF053074 NM_003604 T34235 AF062495 AA353437 AW074506 AA355690 BE225548 AA355309 BE067539 BE067459 AW369656 AW063694 AW063720 AW061615 AW763340 AW674262 AW769443 AW470304 AW769843 BE20102 BE303992 A1220589 AA575429 A1243334 AA654267 AW145177 HW5633 AW62332 AW248441 AW768965 AW735620 AW973202 AA342014 HW5554 AA458783 BE205334 BE262473
	312226	223393_1	AA315703 RW5663 AW769616 AW849575
30	312240	576434_1	R36475 RW2658 AW05300 RW591615 HW1236 RW9817 RW5171 RW3255
	312292	698293_1	AW450103 AW451993 HW1073 A1235250 AW28763
	303820	27484_1	AB037358 AW399417 BE168022 BE207137 AW06125 BE003963 AW965680 AA349466 AA351921 A1482658 BE146202 D31580 AW001199 RW5587 AW372674 AW752776 BE168047 AW964238 AA160849 AW427021 T16898 AA161281 AA143499 A1372573 D38001 AW707135 AA40100 AA145829 AW71235 AW071972 AA315111 AA433705 AA32006 A1231622 AW38354 AA478301 A14183356122 AA178934 AW105270 AA499068 T988383 AA349465 AA353331 D60500 AA15399 AA304088 AW334361 AW339463 AA301780 AW375382 AW262630 AA702282 AA376185 AA064862 AA355373 AA102488 AA100840 AA325211 AA426180 BE362689 HW0452 AA367261 RW4717 AA037180 HW05036 AW095827 AW700441 AA102418 AW859516 AA313505 AW591928 AW062735 AW105602 AW67455 AW56560 AA94146 AW422026 AW392999 AA434926 AA045429 A135488 AW100360 AW428177 A14716207 AB11003 AW026435 RW3554 AW62216 HW05956 AW18217 AW735819 AW31483 HW7542 HW2251 AW359637 H11435 AA156068 AA102485 AW335956 AW034543 BE139052 AW37968 AW075493 AA545107 A1004530 AA548969 AW892221 AW01671 AW570099 AW64590 AA631071 AW770217 AW471322 WB3786 AW134571 AW142199 AW729341 AA807680 AW473618 AA73825 AA605001 RW4038 AW173322 AA622276 AW709620 AA72575 AW670496 RW17334
35	312313	424162_1	AW529341 AW750158 AW21417 AW28557 AW674282 AW194521 AA651778
	312381	742392_1	A1526755 RW5782 RW5791
	312405	765547_1	A1422223 AW24442 MW7603
40	313070	734749_1	AW67164 AW72427 N52323 RW7508
	313156	717700_1	AW01039 AA10047 RW5934
	313179	510895_1	AA327630 AW011663 AW196229
45	311928	261800_1	T62216 AA557503 AW953611 AW024798
	319808	7058_3	T59890 AA609180 AA621130 AW527236 AA431075
	321023	494226_1	AW294516 AA977516 AW075680 AW972654 AW31755 AW64320 AW08443 AA694643 AW389892 AW68770 A2470175 HW26135
50	321024	126702_1	AW724026 BE243567 BE241413 AA624526 AW061158 AW646965 AA524841 AW07859 AW19768 AA604637 AW669821 AW11425 AA694637 AW359501 AW500412 AW614020 A1270184 AA351957
	319587	200730_1	NA6574 AA226737 BE503675 AW161363 HW1831 AA68626 T6946 MW0525 N26680
	312500	83181_1	AW970595 HW3004 AA015335 HW6102 AA516342
55	313250	229754_1	AW805404 A2265537 L51700 AW470054 AW70720 AA327833
	312589	843717_1	AW402461 AW120166 RW8714 D91222
	319928	37979_2	AW820719 AW1273516 AW592957 A1253704 AW31925 RA0866
60	320774	257619_1	AL049443 C01146 AW78278 AW02508
	320664	430084_1	AW160715 AW016237 AA831284 AA749122 AA343680 AA741126 AA946695 AA524841 AW07859 AW19768 AA604637 AW669821 AW11425 AW033130 AW19768 AW243696 AA0846 AW859142 RW2725 AW336730 AW336771 AW350205 AW093231
	312800	412682_1	A1248774 RW73944 HW0813 AA702493
65	312833	404357_1	AA577534 AW06952 HW1242
	320683	234446_1	AA334511 AW505008 BE265237 AW702758 RW50291
	313476	83840_1	AA012020 AA526774 AW3616870 AW625372 AW371245
70	320987	341050_1	N52937 NW3006 NE3029 NW3018 NW3004 RW5876 RW2598 AA525258
	312821	410071_1	AA599325 AW67455 HW75966
	321325	28286_1	AW033130 AW160168 BE260255 AW961689 AW47207 A1347037 AW703994 AA601045 A1559897 AW139033 AW724622 AW172394 AW080700 AA040340 AW796255
75	314146	183043_1	AW827237 AW294348 AW52522 A1305704 AA460256 AW151450 AW04404 AW09484 A155398 A223534 A1015721
	314171	165032_1	AW821896 AW732057 AA400610 AA661559 AA243957 AW186132 AA24048 AW821242 AW820544 AA244056 AA20471
	313552	110315_1	AW69206 AW022652 AA525256 AW018601 AW015865 AW02311 AW026182 AA130656 AW511662 AW070797 RW5185 AA633153 T55272 HW6501 AW070683 AW772654 BE57987 AW517380 T67528 AA426557 AW67155 AW769158 AW347474 HW0698 AW47598 RW0730 BE127732 NW5763 AW38969 AW109671 AW59026 AA169306 AA94412 AA063582 AW168996 AW82576 RW9172 AA434832 AW83041 RW7439 RW6517 RW1140 HW5322 HW4236
80	320771	210125_1	RW4441 HW4256 RW6890 AW75026 AW851774 BE327562 AW00726 AW770471 BE326537 AW218667 AW92356 AA24532 AA301750 AW209012 AA686326 HW059 AW62776
	320779	74700_2	AA615354 AW452966 RW72798 AW659767 AW33067
	320787	600798_1	AW088363 RW3323 AW628638 AA090302
85	312939	264486_1	AA499930 AW470690 RW87831 AA360368 BE166712
	326386	C_K_16	
	326892	CH22_438FG_41_1	CH22_438FG_41_1
90	326684	CH22_4197FG_46_1	CH22_4197FG_46_1
	326721	CH22_4244FG_83_17	CH22_4244FG_83_17
	326838	CH22_8635FG_LINK_EMAC00	CH22_8635FG_LINK_EMAC00
95	326899	A113928	A113928
	326819	CH22_6944FG_LINK_EMAC00	CH22_6944FG_LINK_EMAC00
	326861	CH22_7294FG_LINK_EMAC00	CH22_7294FG_LINK_EMAC00
	326862	CH22_7396FG_LINK_EMAC00	CH22_7396FG_LINK_EMAC00





WO 02/098358

PCT/US02/17594

5

10

15

20

25

30

35

40

45

50

55

60

65

70

75

80

A271939 A7987136 AA812903 AW198919 A113542 A4B43006 C05894 C75127 AW044680 T03756 AW1583349 AA020353 AAB74739  
 A282255 AA010549 AW166991 A5213979 A030490 AE311099 A313396 W61554 A529235 A568189 AA010548 AA478368 A502235  
 A074528 AA382701 W46570 A570312 AA562306 A802069 A452364 AA099398 A4533379 AA010381 AA387721 AA154453 A5023933  
 T33340 A53316 A581354 T11592 A223910 A565343 A565343 AA110559 A561127255 A458259 A5010398 A5010398 AA018659  
 A58907 A A535915 A582782 AA84130 A016181 AA281352 A020492 A001734 AA063581 A5093556 A588413 A5093556 AA01068  
 AW504730 A018658 AW020841 T09170  
 A1130740 A1131523 AW016400 W27900 M81906 AL047197 D69003 N21300 AA34310808 AC310491 AA075434 AAO21104 T06565 W26925  
 A5345393 T78213 AA345626 R19368 R14467 A544936 A3254917 H14308 A2039592 A506898 A506898 A506898 A506898 A506898 A506898  
 A072107 R11220 T51144 R05023 A4572245 AA307771 AW064894 A5064040 A5064040 A5064040 A5064040 A5064040 A5064040 A5064040  
 W22881 W087166 R087166 AW000004 N42531 A5097448 AA180623 T29126 AW651130 AA2242337 A5011735 A687725 A687725 A687725 A687725  
 T7040 A4883195 AW172953 AW072767 A5011932 H14227 A114693 AW40807 A5094262 R40336 A1823887 186837 A795343 A45691  
 A525160 H08422 H09529 H06628 A5033016 B2E45981 A5093509 A453529 N06641 W0895 A506072 T15783 H43128 N08131 A5032599  
 R08117 T01911 T04589 W22253 T102 A4565115 A514737 H45430 AW270027 A3334281 A0371186 A4713687 A4723162 AA010383  
 W22222 A4055114 A11070730 Z44767 F11517 F11518 A5033033 H19191 A687570 A1500448 A4331903 A5052385 A503519 A5062385  
 H05184 A1124826 Z41982 T02068 B0015740 AA481998 A540623  
 N1\_020328 U22970 U22970 A0473634  
 U20514 Z4398 R11803 R53806 A0020316 AW953299 AW493993 AW109942 B6E12986 R52743 R12125 A5052701 N25044 W08328 W18794  
 W01192 AW03981 A5051788 W45261 A5051785 A54370875 B00E1075 A0030310 A0050132 B00E1075 E5074256 AW137481 A477685  
 A565700 A5062228 A46658 A029333 A452366 A4702499 AW195744 A827700 A533051 A1854343 H11624 A4875525 A079228  
 A509592 A9319183 A420739 A504580 A423186 AA132389 A5057832 R08269 A564441 A333978 R19545 A50512718 N09183 A1054206  
 A50511 A5037784 A0805441 A0807239 W1007 A27426 A470999 A621334 T33592 A224482 A459105 A4507448 A452151 A459297  
 B616458 R44735 T15990 A453597 Z29987 188695 A523244 F00027 R03873 T03575 A5091244 A5091244 H03920 A4772510 A462970  
 N93224 A5051255 F00227 H86561 T25741 A1146573 A089149 T30872 H64306 F06653 Z43051 R19934 H11649 H1234 B2E42391  
 A5657794 A5657794 A5657794 A5651458 A5651478 A5651478 A5651478 A5651478 A5651478 A5651478 A5651478 A5651478 A5651478  
 A471167 H03241 AA006339 A4403992 A525801 A445322 AA745088 A087211 H06519 AW189337 A1336215 AA489964 A008918  
 R03275 W087166 A183567 R32706 A450145 A450145 A450145 A450145 A450145 A450145 A450145 A450145 A450145 A450145 A450145  
 A5075945 A4402680 R32395 A5075387 W445970 A4746749 H51720  
 A0797994 N1\_016228 AA404822 T39680 AW166135 A10219254 A10219254 A10219254 A10219254 A10219254 A10219254 A10219254 A10219254  
 A4333387 H12165 F11219 R03274 H72204 AW697408 AW191554 A503936 A505060 A020479 A505060 A4242427 A139304 A502098  
 R40205 AA68044 F08931 A471167 A471167 A471167 A471167 A471167 A471167 A471167 A471167 A471167 A471167 A471167 A471167  
 A0797994 N1\_005479 F0006239 A44021023 B2E44859 A4311615 A4027025 AW096332 A1350040 A359154 A4626865 AW077244  
 AW094407 A424708 A4730059 R78075 AW002924 A961173 A11892003 A0574222 A585446 A435212 A602065 B646128 W06577  
 A023099 A486418 A022750 AW148754 A4509817 A501262 A1374757 B00E46375 B656801 A108264 A000301 A47334 A4828747  
 A5692327 B2E46513 A4232231 A5054100  
 A005200 R70595 A5250728 A505514 A1320255 A505537 N39606 A419721 A4187026 A278952 A825136 A181770 A151140 B2E46117  
 A522210 A5067070 A49605 A505045 A4976383 A5253432 T77725  
 A40558 AW23040 A448663 A4504143 A5058223 A4935374 T19666 T19667 D62239  
 H093805 A4651765 A102726 A102726 A102726 A102726 A102726 A102726 A102726 A102726 A102726 A102726 A102726 A102726  
 A440126 A4012621 A5029432 A501454591 Z29539 A4291617 R0806 A4291783 A4963 A1047408 A1047408 A1047408 A1047408 A1047408  
 B673954 A512425 A1039662 A1044012 A1061735 A0543161 A4743637 A502222 A423022 A507952 A503609 A516571 AW00024  
 A2619954 A431733 A39125 A486306 AW0273910 A7108890 A0009335 A5051913 A5051147 A505294 A107126 A506303 A300860  
 A4011556 A4624327 A535489 B2E19425 A1004356 AW151394 A218466 N66178 A419784 AW242519 A5949907 D60374 A4899283  
 A5087919 A4073400 A4261467  
 A595504 A0117034 A511704 A0151730 A0054543 A51946327 D088702 E080898 E080898  
 A030732 A503898 A4215297 B6E47488 AW17355 A4046224 A4361094 A473328 A50152704 A128330 A307387 A1138263 AA046116  
 A1219814 A315431 A01169599 A4842005 A5028002 A5028002 A4401340 A413781 AA292833 A031271 A0227867 R31381 A502332 A1896186  
 AW46716 A4492824 A1483091 A453081 A4450417 A4568208 AW007397 A5057282 A4628439 AW1206551 A47167351 A4872754  
 A071619 B2E27791 A5011828 A4744724 A4027816 A15131708 A4031641 A5032786 A4717401 A765196 AW08076 A5073208 A4567164  
 A4744556 A4888510 A572276  
 A4062977 A082955 A406296  
 W975991 A463160 A144256708 AW139028 A061816 A112247  
 H01809 A410066 R27719 R77038 T23782 H140151 N404538 R506233 T94741 A1110626 F053500 W35141 A236329 H15136 AW143456  
 N23362 AA682672 H03110 A11168530 N32246 T94740 A4236651 A236235 N23655 R77220 N16114 A4814426 A080487 A509474 A400155  
 N30798 A431754 A61763 A1272814 A322944 A4193683 A1183993 A1183991 Z39970 F04881 A893457 N26707 B6E35336 R92734  
 A494599  
 A533652 A5076543 A800374 A48793 AW195220 A898934 A565252 C01574 R62583 B6463212  
 N66563 R02024  
 B600767 R06103 N68542 B600791  
 A1945841 A409278 N70242  
 A503737 R0437 R54797  
 R44892 A89855 A900018 A4878965 A479353 A453159 A424114 A095174 A897714 T32446 T16534 A4028917 D63033 A5059291  
 A4067730 A4135210 A5081959 A4457804  
 A162546 A1027178 N11\_004362 A106570 A077057 A000336 A5057063 A5085365 A8004900 A4049023 A396069 A0010001 T11719  
 R11545 A219450 AW172082 A5092466 A5094002 A3375131 A677921 A002524 A355330 B6E459771 A2120096 D06574 A0213327  
 A4252130 R19271 A522192 A4252079 A1488677 A1608022 A4911675 R25458 A1021796 A411675 A505051 A5024718 A1024557 A393640  
 A020220 A50155139 R26283 A240607 F44991  
 A1761104 A14356 A802687 A007720 A007720 A007720 A007720 A007720 A007720 A007720 A007720 A007720 A007720 A007720 A007720  
 A118425 A0273879 A202140 A4255824 A0701863 A4233958 A4252086 A0909101 A4074094 A0020951 A0020951 A0020951 A0020951 A0020951  
 AA303374 A4256521 A4565336 A5197149 A1120920 A4006028 A4720991 A4236827 A4730777 A515110 A118075 A4801939 A4677413  
 W33115 A4822549 A4249038  
 A4267450  
 N74302 N74444 A4287852 A4206513 A4287882 W02641 W03165  
 B6541042 A0069754 AW0684160 AW189935 R27296 R27297 H0475 R8704 A587583 R78627 A421458 W02221 AA303125 A805269  
 R45578 A5715977  
 A465308 A221943 A27578 A425637 A4406577 A4479917  
 A5082529 A5065054 A538451 N1\_006302 A5316715 A10629 A5069300 AA314225 F007091 A4421527 B6072059 A810703  
 AW194116 AW192795 A0075324 A258837 A634717 A080057 A5151674 A080254 AW190656 A5084247 A080040 AW15248 A5002338  
 A501474 A445913 A528325 A453742 A438796 A4089495 A735767 A5034001 A457938 A5030845 A4582017 D0759935 E024320  
 A5178847 A4071069 A212001 A509617 A0808914 A0227220 A801054 A4532861 A5073291 A51580301 A431540 A528261 A8011784  
 A5095627 A502366 A477690 A509391 A080443 A5040451 A505949 A5057981 A4849431 A422882 A507851 A21154943 A5116874



WO 02/098358

PCT/US02/17594

AV021536 AV1118330 AA515358 D56610 AA494082 D66934 T97774 AA473546 R74350 R84834 AA579200 D56616 C03207 D57301 N52416  
D56928 R79209 D56925 AA020879 D45546 AB58769 R20750 T09381 F01435 AW627906 D58202 AI033993 F01912 H27552 AA174191  
T15515 AW023216 AA434146 H83387 AI346751 V01512 V01512 AA576407 AW365140 AA937471 BE174681 AI568820 D274663 R65530  
AI148426 H83398 AW79534  
5 332577 89088\_2 AI826589 AV248872 H69511 AT48806 AW779557 AI062254 AI890377 AW151271 AI336374 AK34503 AA777065 AI590131 H37767 AI895058  
H69512 AA045480 N27343 AI573008 AW130925 AI632838 AW594603 AW000790 AI208239 AI275835 AW1000294 AN021587 AW273456  
AA650725 AW469424 AI400222 AI025723 BE046146 AI128968 BE350462 AW302951 AI299977 AA284809 A1840358 AW470364 A1241754  
AI650048 AW090327 H16377 AW615318 D66021 AI934336 AW118536 AJ041281 AA614238 R6918 AW571741 AV1516952 AV1572322  
AV151518 AI708345 AI329522 Z4518 AI89930 AA469975 AB57819 AI810684 AT070744 AI370410 BE385985 Z4476 BE302481 BE02532  
AA456755 AA4159 D60022 Z14004 AA021099 AA284872 BE266647 AW249292  
10 332640 4172\_1 BE568452 BE297396 AA440593 AW732480 AW069736 BE548067 AA207229 AF044588 NM D03981 BE268954 AW444578 AA471151  
BE290747 AW732555 AA074582 BE336956 AW008764 AA191159 BE092129 AA310614 AW595867 AA312276 AW750027 AW750046  
AW750032 AW750024 AA108983 AW750054 AA048409 AW750030 BE151675 AA478359 N65721 AA105614 I70079 H75580 BE25401  
AA454518 AI3027263 AA625405 AA417152 AA004290 AA57354 AW863151 AW863151 AA702179 AB24143 AW57185 BE005198 AA190630  
15 4172\_1 AI638795 AI609113 AI056239 BE537023 BE464668 AA634413 BE208066 BE208833 AW2520803 AI337375 AA478510 BE501624 AI814763  
AV1594726 AI31408 AA827285 AA189108 AW694169 BE618589 BE618040 AI130398 AA632206 A006126 AB638100 AA725439 AI379107  
AI288872 H14601 AI678151 AI263619 AI552121 AB679222 H61249 AA552246 AA470300 AI86543 AI534494 AI038187 AA176564 AA573241  
AI734525 AA043507 BE207965 AI291838 N73385 N73535 AW859561 AA808510 AB699813 AW166044 AV104716 H05808 AA248270  
20 332732 5486\_1 BE538022 N56013 AA621586 AA445727 D19671 AV192090 N54283 H73339 AA109689 BE273424 BE560082 AW569012 AA313652  
AV750034 BE072537 BE207947 AW732361 AA440336 D29574  
AF191819 NVA\_015516 BE546494 AL110276 R13844 BE313596 BE336912 R19704 R18703 AA458968 T70552 BE336911 T60307 BE149749  
BE271848 BE271802 AA409229 Z45402 T04900 AA302745 AA004651 T95706 H14807 AI399901 C03221 T72431 AA0471185 AI3433257  
A269100 BA345077 AW965160 H27581 R49910 H25380 AA335081 AV973283 T79590 AW169447 T64172 AI744557 AI342356 AA336102  
25 332732 5486\_1 AA335299 BE298375 AI140634 AA088181 AI860314 AT738613 T70902 R42077 A884568 AA469798 AI130828 AA009735 H25381 AW612425  
RA8801 H27507 H30105 H44671 AR31362 AA338470 AV0144412 AA520519 AA049801 AV1599435 AI039657 H14614 AA974256 R42078  
AA245750 T61696 A555202 AB04139 AB07313 AI041404 AA437136 AB13332 AI147891 AA457945 AV1197727 AI274309 AT758036 AI59048  
NA972077 M83390 R36989 R71936 AB67482 T40081 Z41115 AA772775 T41013 AB656891 T40996 AB26822 N59464 AW955524 AA088651

WO 02/098358

PCT/US02/17594

TABLE 1C

Play: Unique number corresponding to an Eos probe

Ref: Sequence source. The 7 digit numbers in this column are Genbank Identifier (GI) numbers. "Dunham I. et al." refers to the publication entitled "The DNA sequence of human chromosome 22." Dunham I. et al. (1995) *Human Genome* 402:489-495.

Strand: Indicates DNA strand from which probe was predicted.

NI\_position: Indicates nucleotide positions of predicted exons

Play	Ref	Strand	NI_position
5	332792	Dunham, I. et al.	Plus 73381-73788
	333135	Dunham, I. et al.	Plus 3361209-3361369
	333137	Dunham, I. et al.	Plus 3367643-3367726
	333138	Dunham, I. et al.	Plus 3369205-3369323
	333139	Dunham, I. et al.	Plus 3369485-3369571
15	333516	Dunham, I. et al.	Plus 5570204-5570380
	333517	Dunham, I. et al.	Plus 5570729-5570925
	333795	Dunham, I. et al.	Plus 7807689-7807756
	333796	Dunham, I. et al.	Plus 7808233-7808319
	333808	Dunham, I. et al.	Plus 7880600-7880775
20	333809	Dunham, I. et al.	Plus 7880600-7880775
	333845	Dunham, I. et al.	Plus 8006232-8006945
	333849	Dunham, I. et al.	Plus 8018323-8018472
	334101	Dunham, I. et al.	Plus 9973419-9973550
	334518	Dunham, I. et al.	Plus 15176123-15176470
25	334891	Dunham, I. et al.	Plus 19299770-19299944
	334899	Dunham, I. et al.	Plus 19315198-19315311
	334900	Dunham, I. et al.	Plus 19315678-19315743
	334902	Dunham, I. et al.	Plus 19317053-19317155
	334905	Dunham, I. et al.	Plus 19322553-19322680
30	334906	Dunham, I. et al.	Plus 19323493-19323590
	335044	Dunham, I. et al.	Plus 26842395-26842882
	335149	Dunham, I. et al.	Plus 27407441-27407587
	335809	Dunham, I. et al.	Plus 26310772-26310909
	335810	Dunham, I. et al.	Plus 26314767-26314846
35	335824	Dunham, I. et al.	Plus 26376959-26376942
	336864	Dunham, I. et al.	Plus 29161595-29161837
	336721	Dunham, I. et al.	Plus 3371523-3371598
	337182	Dunham, I. et al.	Plus 23934889-23934962
40	337674	Dunham, I. et al.	Plus 3332616-3332697
	337675	Dunham, I. et al.	Plus 3335398-3335505
	337755	Dunham, I. et al.	Plus 3971764-3971800
	338038	Dunham, I. et al.	Plus 6138219-6138392
	338316	Dunham, I. et al.	Plus 17089711-17089886
45	338124	Dunham, I. et al.	Minus 3318017-3317932
	333743	Dunham, I. et al.	Minus 7572161-7573080
	334221	Dunham, I. et al.	Minus 12730944-12730387
	334222	Dunham, I. et al.	Minus 12732417-12732289
	334282	Dunham, I. et al.	Minus 13285293-13285178
	334502	Dunham, I. et al.	Minus 14488935-14488826
50	334578	Dunham, I. et al.	Minus 15004452-15004304
	334951	Dunham, I. et al.	Minus 20147708-20147502
	335289	Dunham, I. et al.	Minus 22306393-22306706
	335290	Dunham, I. et al.	Minus 22309253-22308811
	335293	Dunham, I. et al.	Minus 22314408-22314575
55	335682	Dunham, I. et al.	Minus 25421215-25421093
	335753	Dunham, I. et al.	Minus 25761515-25761444
	335755	Dunham, I. et al.	Minus 25783806-25783747
	335796	Dunham, I. et al.	Minus 25764339-25764291
	336962	Dunham, I. et al.	Minus 21680660-2167993
60	336884	Dunham, I. et al.	Minus 21680660-2167993
	337603	Dunham, I. et al.	Minus 1299296-1299194
	338581	Dunham, I. et al.	Minus 22311965-22311896
	338682	Dunham, I. et al.	Minus 22312594-22312465
	338186	Dunham, I. et al.	Minus 32339211-32339097
65	326889	5867067	Plus 2232829-2232891
	330032	6652596	Plus 85177-85237
	330933	6652596	Plus 86653-86723
	326213	5867224	Minus 60751-60927
	326816	6552458	Plus 198354-198438
70	327110	6117842	Plus 94608-94785
	327821	5867968	Plus 131080-1311232
	328164	5868068	Minus 27080-27226
	328648	6004473	Plus 424879-424959
75	329365	5868938	Minus 107687-107765

Table 2A lists about 1165 genes selected to have an interesting expression pattern during androgen withdrawal of prostate cancer tissue. These genes were selected by analysis of variance, such that the P value is less than 0.01, the 90th percentile exhibits a minimum of 100 average intensity across all samples, and a comparison of any group means shows a minimum 3 fold change. The interesting expression patterns can be broadly defined into the following categories:

1. Genes that are expressed early in the time course of androgen withdrawal, then drop off in expression, and then express again with emergence of androgen-independence (hi-to-lo pattern in Table 2A).
2. Genes that are expressed early in the time course, then drop off in expression immediately after androgen-withdrawal, and do not express again with emergence of androgen-independence (hi-to-lo pattern in Table 2A).
3. Genes that are expressed early in the time course, then drop off in expression after several days of androgen withdrawal, and do not express again with emergence of androgen-independence (hi-to-lo pattern in Table 2A).
4. Genes that are not expressed early in the time course, but express only with emergence of androgen-independence (lo-to-hi pattern in Table 2A).
5. Genes that are not expressed early in the time course, but then express as androgen is withdrawn and continue to express with emergence of androgen-independence (lo-to-hi pattern in Table 2A).
6. Genes that are not expressed early in the time course, but then express as androgen is withdrawn and drop off again with emergence of androgen-independence (lo-to-hi pattern in Table 2A).

Table 2B lists accession numbers for primers lacking a uniqueID in Table 2A. For each primer set is listed a gene cluster number from which oligonucleotides were designed. Gene clusters were compiled using sequences derived from Genbank ESTs and mRNAs. These sequences were clustered based on sequence similarity using Clustering and Alignment Tools (Doubio/Tal, Oakland California). Genbank accession numbers for sequences comprising each cluster are listed in the "Accession" column.

Table 2C lists genomic positioning for primers lacking unique IDs and accession numbers in Table 2A. For each predicted exon is listed genomic sequence source used for prediction. Nucleotide locations of each predicted exon are also listed.

TABLE 2A: ABOUT 1165 GENES SELECTED TO HAVE AN INTERESTING EXPRESSION PATTERN DURING ANDROGEN WITHDRAWAL OF PROSTATE CANCER TISSUE

Key: Unique Eas primer identifier number

ExAccn: Exemplar Accession number, Genbank accession number

UniqueID: Unique number

Unique Title: Unique gene title

Pattern: Broadly defined expression patterns during androgen withdrawal

Key	ExAccn	UniqueID	Unique Title	Pattern
433412	AB53729	Hs.6165	CGI-44 protein, sulfide dehydrogenase II	lo-to-hi-to
429597	AK001270	Hs.196086	hypothetical protein FLJ10465	lo-to-hi-to
442751	AB68167	Hs.131044	ESTs	lo-to-hi-to
420820	U26096	Hs.336535	Homo sapiens, clone IMAGE41794E2, mRNA	lo-to-hi-to
422267	AB535044	Hs.114012	KIAA1216 protein	lo-to-hi-to
416953	NS1537	Hs.29046	ESTs	lo-to-hi-to
413277	I24177	Hs.75262	cathepsin O	lo-to-hi-to
410209	AB53661	Hs.60548	hypothetical protein PR01535	lo-to-hi-to
428623	AW974540	Hs.96526	ESTs	lo-to-hi-to
435847	V8321	Hs.35760	CD4017 protein	lo-to-hi-to
443967	AW294013	Hs.209842	ESTs	lo-to-hi-to
440838	AA907075	Hs.131307	ESTs	lo-to-hi-to
440254			Target Exon	lo-to-hi-to
431697	H67740	Hs.36540	ESTs, Weakly similar to ALUM_HUMAN ALU S	lo-to-hi-to
432114	AL59021	Hs.8934	ESTs	lo-to-hi-to
446387	AW275603	Hs.200712	ESTs	lo-to-hi-to
414094	H16088	Hs.31433	ESTs	lo-to-hi-to
424005	AB535041	Hs.137507	varg (van gogh, Drosophila)-like 2	lo-to-hi-to
424421	H7722	Hs.162661	death effector domain-containing	lo-to-hi-to
449749	AB58611	Hs.49750	ESTs	lo-to-hi-to
463686	BE604731	Hs.138827	ESTs	lo-to-hi-to
472721	L15409	Hs.174007	von Hippel-Lindau syndrome	lo-to-hi-to
427115	AJ247162	Hs.209483	ESTs, Weakly similar to KIAA1074 protein	lo-to-hi-to
425980	AA366661		gbcEST77963 Pancreas tumor Hs. Homo sapi	lo-to-hi-to
412482	AW962604		gbcEST734677 MAGE resequences, MAGG Homo	lo-to-hi-to
438982	AA827896		glzod560c2a1 NCL_GCAP, GCB1 Homo sapiens	lo-to-hi-to
423473	U94780	Hs.117242	meningioma expressed antigen 6 (cell-to-c	lo-to-hi-to
404211			NW_005555 Homo sapiens myosidin/lymphoid	lo-to-hi-to
423019	AB40185	Hs.283626	ESTs	lo-to-hi-to
443559	AB07865	Hs.289899	ESTs, Moderately similar to ALUS_HUMAN A	lo-to-hi-to
444291	AB596022	Hs.193589	TAR DNA binding protein	lo-to-hi-to
428065	AB54046	Hs.167313	ESTs	lo-to-hi-to
442566	R37337	Hs.12111	ESTs	lo-to-hi-to
442202	BC272862	Hs.106534	hypothetical protein FLJ22625	lo-to-hi-to
439456	AT52409	Hs.109314	hypothetical protein FLJ20980	lo-to-hi-to
423475	AL055533		Homo DNA sequence from clone RPS-1046G31	lo-to-hi-to
479522	D63209	Hs.5944	solute carrier family 11 (proton-coupled	lo-to-hi-to
451987	AA815082	Hs.77554	Homo sapiens cDNA FLJ14667, iso, clone TH	lo-to-hi-to
453408	AB04732	Hs.256953	ESTs	lo-to-hi-to
444004	N26842	Hs.301444	KAA1573	lo-to-hi-to
452991	AA156442	Hs.192519	KAA1600 protein	lo-to-hi-to
434855	AW050449	Hs.116507	ESTs	lo-to-hi-to
440819	AB09444	Hs.202108	ESTs	lo-to-hi-to
419626	AB21895	Hs.193481	ESTs	lo-to-hi-to
422072	AB218255	Hs.111138	KAA40712 gene product	lo-to-hi-to
453459	BE047032	Hs.257789	ESTs	lo-to-hi-to
419038	AW134924	Hs.190325	ESTs	lo-to-hi-to
413243	AA789296	Hs.193687	ESTs	lo-to-hi-to
432079	AW972748		gbcEST384640 MAGE resequences, MAGL Homo	lo-to-hi-to

WO 02/098358

PCT/US02/17594

	441328	A1982794	Hs.159473	ESTs	to-to-hi-to
	416508	R39769	Hs.206132	ESTs, Moderately similar to ALU8_HUMAN A	to-to-hi-to
	451066	A175860	Hs.206132	ESTs	to-to-hi-to
5	446017	N85238	Hs.55165	ESTs	to-to-hi-to
	447104	R19085	Hs.210706	Homo sapiens cDNA FLJ13182 fls, clone NT	to-to-hi-to
	447211	AL161961	Hs.17767	KIA1554 protein	to-to-hi-to
	447765	AJ014112	Hs.161390	ESTs	to-to-hi-to
	425940	N85776	Hs.161390	gb:EST22297 fcdl brain, Striatagene (cat	to-to-hi-to
10	443141	A1140497	Hs.161390	gb:aw16200.s1 Soares, fetal liver, spleen,	to-to-hi-to
	414555	N65569	Hs.76422	phospholipase A2, group IIA (platelets,	to-to-hi-to
	432677	NM_004462	Hs.278611	UDP-N-acetyl-alpha-D-glucosamine:poly	to-to-hi-to
	423091	A506539	Hs.97927	ESTs	to-to-hi-to
	423026	H80546	Hs.101590	gb:U86022.11 Soares fetal liver spleen	to-to-hi-to
15	444040	AF204231	Hs.182982	ccl3n-67	to-to-hi-to
	441111	A306667	Hs.126594	ESTs	to-to-hi-to
	418538	AW385224	Hs.35198	ectonucleotide pyrophosphatase/phosphodi	to-to-hi-to
	415989	AA172179	Hs.294022	ESTs	to-to-hi-to
	429615	A7266021	Hs.211562	ATP-binding cassette, sub-family A (ABC1	to-to-hi-to
20	427774	AA27683	Hs.180737	Homo sapiens clone Z3664 and Z3905 mRNA	to-to-hi-to
	438585	AA811371	Hs.123362	ESTs	to-to-hi-to
	424776	A067931	Hs.164595	ESTs	to-to-hi-to
	413786	AJ013780	Hs.13500	ESTs	to-to-hi-to
	421077	A3000061	Hs.101590	hypothetical protein	to-to-hi-to
25	448837	A0261700	Hs.145544	ESTs	to-to-hi-to
	442822	AL046256	Hs.23437	Homo sapiens cDNA FLJ13535 fls, clone PL	to-to-hi-to
	414065	AJ0151573	Hs.271249	Homo sapiens cDNA FLJ13380 fls, clone PL	to-to-hi-to
	432527	AJ075028	Hs.102754	ESTs	to-to-hi-to
	412093	BE242691	Hs.14947	ESTs	to-to-hi-to
30	417211	A743770	Hs.160513	ESTs, Weakly similar to KIA00522 protein	to-to-hi-to
	417280	AJ173116	Hs.250103	ESTs	to-to-hi-to
	422445	A302438	Hs.20596	Homo sapiens mRNA from chromosome 5q21-2	to-to-hi-to
	438624	AA889055	Hs.123466	ESTs	to-to-hi-to
	442343	AA952480	Hs.129874	ESTs	to-to-hi-to
35	401416			C14005339.g17459502.g174665 outer	to-to-hi-to
	417176	AJ176906	Hs.42346	calcitonin-binding protein calcitonin-1	to-to-hi-to
	416663	A872360	Hs.209293	ESTs	to-to-hi-to
	448295	AJ137268	Hs.270954	ESTs	to-to-hi-to
	428948	AJ72631	Hs.36190	ESTs	to-to-hi-to
40	445467	A229832	Hs.15617	ESTs, Weakly similar to ALU4_HUMAN ALU S	to-to-hi-to
	419562	A010108	Hs.151500	ESTs	to-to-hi-to
	416239	AL038450	Hs.46948	ESTs	to-to-hi-to
	428024	A948688	Hs.266619	ESTs	to-to-hi-to
	433264	AJ079470	Hs.96649	Homo sapiens cDNA FLJ11492 fs, clone HE	to-to-hi-to
45	424332	AA338919	Hs.101615	ESTs	to-to-hi-to
	442369	A165071	Hs.159983	ESTs	to-to-hi-to
	420717	AA284447	Hs.271887	ESTs	to-to-hi-to
	438684	AA536114	Hs.221612	ESTs	to-to-hi-to
	440260	AJ72987	Hs.7130	caprin IV	to-to-hi-to
50	426289	H15302	Hs.168950	Homo sapiens mRNA: cDNA DKFZp666A1046 f	to-to-hi-to
	428396	A249368	Hs.98558	ESTs	to-to-hi-to
	407276	A951118	Hs.326126	Homo sapiens breast cancer antigen NY-ER	to-to-hi-to
	429389	A020266	Hs.54037	ectonucleotide pyrophosphatase/phosphodi	to-to-hi-to
	442150	A368158	Hs.70983	PTPL1-associated RhoGAP 1	to-to-hi-to
55	415787	H01463	Hs.93534	ESTs	to-to-hi-to
	430685	A090234	Hs.191666	ESTs, Weakly similar to GNMSLL reovirus	to-to-hi-to
	443794	N04154	Hs.26290	ESTs	to-to-hi-to
	446215	AJ0621329	Hs.14368	SH3 domain binding glutamic acid-rich pr	to-to-hi-to
	441285	NM_002374	Hs.167	microtubule-associated protein 2	to-to-hi-to
60	447838	BE614081		gb:6015381F1 NIH_MGC_71 Homo sapiens c	to-to-hi-to
	435746			BNP000100226812-KIA1494 protein (Frag	to-to-hi-to
	434022	R18374	Hs.117956	ESTs	to-to-hi-to
	435714	A0699325	Hs.269880	ESTs	to-to-hi-to
	439848	AJ0979249		gb:EST391359 MAGI2 resequences, MAGP Homo	to-to-hi-to
65	421974	AA301270		gb:EST14152 Testis tumor Homo sapiens cd	to-to-hi-to
	433522	A1817347	Hs.44488	Homo sapiens clone TCCCTA00151 mRNA sequ	to-to-hi-to
	449919	A674685	Hs.200141	ESTs	to-to-hi-to
	407192	AA009200		gb:af126a2.s1 Soares_testis NIH Homo sap	to-to-hi-to
	436169	AA886311	Hs.17602	Homo sapiens cDNA FLJ1281 fls, clone MA	to-to-hi-to
70	418624	A774080	Hs.104211	ESTs	to-to-hi-to
	432432	AA541323	Hs.115831	ESTs	to-to-hi-to
	426172	AA371307	Hs.125056	ESTs	to-to-hi-to
	401093			C12000586.g16330167.g16330167 (A	to-to-hi-to
	425716	NM_006379	Hs.171921	same domain, Immunoglobulin domain (Ig)	to-to-hi-to
75	439669	AJ002186	Hs.222399	CESP1 protein	to-to-hi-to
	461720	AJ097085	Hs.200653	ESTs	to-to-hi-to
	428163	AA884766		gb:am20a10.s1 Soares_NFL_T_GBC_S1 Homo s	to-to-hi-to
	432435	BE216886	Hs.282070	ESTs	to-to-hi-to
	406170	AJ030416	Hs.51835	ESTs	to-to-hi-to
	433530	BE349534	Hs.281789	ESTs	to-to-hi-to
80	425776	U25126	Hs.159499	parathyroid hormone receptor 2	to-to-hi-to
	430068	AA464964		gb:z88010.s1 Soares ovary tumor N6HOT H	to-to-hi-to
	427225	AA315703	Hs.199993	ESTs, Weakly similar to ALU6_HUMAN H	to-to-hi-to







WO 02/098358

PCT/US02/17594

	400111		Eosin Control	to-to-hi-hi
	405445	A015709	Homo sapiens mRNA; cDNA DKFZp596I0222 (F	to-to-hi-hi
	401953		C15001262:g[7304881]e[PNP_038528.1]ca	to-to-hi-hi
5	402785		C1500867:g[1733663]e[PKP_011474.1]C	to-to-hi-hi
	425454	AA579658	KAA1457 protein	to-to-hi-hi
	414343	AL036166	coated vesicle membrane protein	to-to-hi-hi
	421970	AF227156	RNA polymerase I transcription factor RR	to-to-hi-hi
	422592	BC081857	root (required for cell differentiation,	to-to-hi-hi
10	413431	AW045428	ubiquitin-conjugating enzyme E2N (homolo	to-to-hi-hi
	425745	J03526	uridine monophosphate synthetase (portal	to-to-hi-hi
	402327		NM_001087*Homo sapiens angio-associated	to-to-hi-hi
	402332		Target Exon	to-to-hi-hi
	402395		Target Exon	to-to-hi-hi
15	459549	AW283364	ESTs	to-to-hi-hi
	401512		NM_014080*Homo sapiens dual oxidase-like	to-to-hi-hi
	448522	AL046505	ESTs, Weakly similar to STK2_HUMAN SERIN	to-to-hi-hi
	402501		ENSP00000251912*HAA1617 protein (Fragm	to-to-hi-hi
20	452324	W81486	ESTs	to-to-hi-hi
	453145	AC039952	ESTs	to-to-hi-hi
	304645	AW82432	ESTs	to-to-hi-hi
	401730		NM_012448*Homo sapiens signal transduce	to-to-hi-hi
	435265	T03890	ESTs, Highly similar to AFX MOUSE HOMEOB	to-to-hi-hi
	400375	NM_014115	NM_014115*Homo sapiens PRO0113 protein	to-to-hi-hi
25	412151	AA100529	Homo sapiens cDNA: FLJ23190 fls, clone L	to-to-hi-hi
	410498	AA355749	gbc:EST54459, Arkalat T-cells V1 Homo sapie	to-to-hi-hi
	405044		NM_014630*Homo sapiens UAA02111 gene pr	to-to-hi-hi
	413159	AW161061	ESTs, Weakly similar to zlac finger prot	to-to-hi-hi
	402101		ENSP00000217725*Laminin alpha-1 chain p	to-to-hi-hi
30	455019	AW850818	gblt3-CT0220-091159-026-A03 CT0220 Homo	to-to-hi-hi
	446826	AK000265	hypothetical protein FLJ20619	to-to-hi-hi
	412180	AW999795	gblt3-HA0075-130403-332-06 NM0075 Homo	to-to-hi-hi
	407273	AJ132960	gblt3-Homo sapiens mRNA for immunoglobulin	to-to-hi-hi
	452895	BE589229	phosphoenolpyruvate kinase	to-to-hi-hi
35	415117	H19460	ESTs	to-to-hi-hi
	421934	A792302	potassium inwardly-rectifying channel, s	to-to-hi-hi
	416309	R84594	cAMP responsive element binding protein	to-to-hi-hi
	444578	T80795	ESTs	to-to-hi-hi
	401966		C17000574:g[8523190]e[PNP_060176.1]ty	to-to-hi-hi
40	444550	AW444882	ESTs	to-to-hi-hi
	403885		Target Exon	to-to-hi-hi
	405435		Target Exon	to-to-hi-hi
	422594	C05003	hypothetical protein FLJ12847	to-to-hi-hi
	422512	AW405973	ESTs	to-to-hi-hi
45	412748	BE083158	Homo sapiens cDNA: FLJ23313 fls, clone H	to-to-hi-hi
	403704		Target Exon	to-to-hi-hi
	440507	H06994	gbc:Y81607.z1 Soares Infant brain T181 H	to-to-hi-hi
	405503		C7000539:gblt2801Zp[JA53933] myosin I	to-to-hi-hi
	459123	R00602	gbc:Y4041.1 Soares Infant liver spleen	to-to-hi-hi
50	454261	AF216077	Homo sapiens clone HB-2 mRNA sequence	to-to-hi-hi
	458956	BE220675	gblt:H811.x1 NCI_CGAP_Lu24 Homo sapiens	to-to-hi-hi
	418367	AA326035	hypothetical protein DKFZp344L0716	to-to-hi-hi
	444553	A167530	ESTs	to-to-hi-hi
	405811		NM_024610*Homo sapiens hypothetical prot	to-to-hi-hi
55	429461	A1188219	ESTs, Weakly similar to HSG2_HUMAN DNAX	to-to-hi-hi
	423378	BE313601	hypothetical protein FLJ22558	to-to-hi-hi
	458516	BE010749	ESTs	to-to-hi-hi
	404039		ENSP00000247650*Hypothetical 177.6 kDa	to-to-hi-hi
	454148	AW732837	nsopharyngeal carcinoma susceptibility	to-to-hi-hi
60	412878	AA115575	ESTs	to-to-hi-hi
	446298	A911333	ESTs	to-to-hi-hi
	405925		NM_002439*Homo sapiens mtuS (E. coli) h	to-to-hi-hi
	424576	BE154142	ESTs	to-to-hi-hi
	451601	N92100	centricornal protein 1	to-to-hi-hi
65	422395	AA310177	DKFZP434B0335 protein	to-to-hi-hi
	434333	AA166153	stromal cell protein	to-to-hi-hi
	413509	BE145419	gblt:LS4T0196-291099-009-E01 HT0196 Homo	to-to-hi-hi
	419504	A068585	ESTs	to-to-hi-hi
	446886	AF289120	CGI-204 protein	to-to-hi-hi
70	401209		C12000516:g[717004]e[PNP_057914.1]N1	to-to-hi-hi
	423554	MC0516	glutamine-fructose-6-phosphate transamin	to-to-hi-hi
	439803	AA001021	thyroid hormone receptor interactor 8	to-to-hi-hi
	424993	AA349729	gblt:EST149730 Gall bladder 1 Homo sapiens	to-to-hi-hi
	406122	AA320552	hypothetical protein FLJ10718	to-to-hi-hi
75	409498	NM_001523	hydraman synthesis 1	to-to-hi-hi
	408214	AL120445	hypothetical protein FLJ21943	to-to-hi-hi
	421911	AL041520	gblt:DKFZp434G2317.z1_434 (synonym: htes3)	to-to-hi-hi
	407813	AL122247	KAA0872 protein	to-to-hi-hi
	425211	H18697	proguinidin (proguinogen C)	to-to-hi-hi
	427772	AW050380	Homo sapiens clone 24416 mRNA sequence	to-to-hi-hi
80	419733	AW362955	Homo sapiens cDNA FLJ14415 fls, clone HE	to-to-hi-hi
	428260	AW290896	ESTs, Weakly similar to S65657 alpha-1C-	to-to-hi-hi
	427063	NM_006353	Sec23 (S. cerevisiae) homolog 3	to-to-hi-hi

WO 02/098358

PCT/US02/17594

5	413583	AA604379	He.86211	hypothetical protein	to-to-hi-hi
	407355	AA846203	He.153974	ESTs, Weakly similar to ALU1_HUMAN ALU S	to-to-hi-hi
	454003	AA058944	He.116602	Homo sapiens, clone IMAGE4164008, mRNA,	to-to-hi-hi
	425522	AL363630	He.155637	protein kinase, CNA-activated, catalytic	to-to-hi-hi
	402240			Target Exon	to-to-hi-hi
	421857	AA481078	He.109045	hypothetical protein FLJ10408	to-to-hi-hi
	408603	R25293	He.326416	Homo sapiens mRNA, cDNA DKFZp566H1916 (f	to-to-hi-hi
	437389	AL315957	He.271595	hypothetical protein DKFZp62M15	to-to-hi-hi
10	427146	AF091036	He.184627	KIAA0115 protein	to-to-hi-hi
	402277			Ess. Control	to-to-hi-hi
	403995			C110002957g1127372837g10XP_012163.1	to-to-hi-hi
	403818			Target Exon	to-to-hi-hi
	402758			C110018997g112726337g10XP_010672.1	to-to-hi-hi
	403708			Target Exon	to-to-hi-hi
15	405610			ENSP000002410655*CDNA	to-to-hi-hi
	414742	AA749230	He.28403	ducl-1-yl-phosphate (UDP-N-acetylglucosam	to-to-hi-hi
	420757	X78592	He.99915	androgen receptor (Rhythoblastadione 1	to-to-hi-hi
	400965			C110021907g1127372837g10XP_012163.1	to-to-hi-hi
	401192			Target Exon	to-to-hi-hi
20	404407			Target Exon	to-to-hi-hi
	401405			Target Exon	to-to-hi-hi
	403055			C20022107g1127372837g10XP_006682.2	to-to-hi-hi
	404681			C20022107g1127372837g10XP_006682.2	to-to-hi-hi
25	433827	AF078886	He.284296	Homo sapiens cDNA, FLJ22993, tis, clone K	to-to-hi-hi
	410204	AJ243425	He.326035	early growth response 1	to-to-hi-hi
	423842	BE297635	He.3069	head neck T202 protein 9B (mortalin-2)	to-to-hi-hi
	401979			Target Exon	to-to-hi-hi
	433990	AA137152	He.280049	phosphoserine aminotransferase	to-to-hi-hi
30	403725			Target Exon	to-to-hi-hi
	413587	AA156164	He.286241	protein kinase, cAMP-dependent, regulato	to-to-hi-hi
	422614	AJ908006	He.296352	Homo sapiens cDNA, FLJ14455, tis, clone HE	to-to-hi-hi
	400275			NM_006513*Homo sapiens seryl-tRNA synth	to-to-hi-hi
	402810			NM_004930*Homo sapiens capping protein	to-to-hi-hi
35	425249	BE266289	He.27693	peptidylprolyl isomerase (cyclophilin) 4	to-to-hi-hi
	445677	H85571	He.6838	ras homolog gene family, member E	to-to-hi-hi
	428770	AK001667	He.193128	hypothetical protein FLJ13805	to-to-hi-hi
	428403	AJ393046	He.326159	leucine rich repeat (in FLJ) interact	to-to-hi-hi
	434647	W74158	He.103189	lipopolysaccharide specific response-68	to-to-hi-hi
	402807			ENSP000000322293*CDNA	to-to-hi-hi
40	413992	V26276	He.133075	RNA, U2 small nuclear	to-to-hi-hi
	407191	AA608751		glucosylated, a1 Stralagone lung carcinoma	to-to-hi-hi
	403328			Target Exon	to-to-hi-hi
	411984	NM_005419	He.72988	signal transducer and activator of trans	to-to-hi-hi
45	451017	BE391847	He.181173	hypothetical protein MGIC10771	to-to-hi-hi
	404108			C70009117g11235142g10AA14470.1 (ACO	to-to-hi-hi
	407819	R42185	He.102220	ESTs	to-to-hi-hi
	435876	AW512586	He.160271	G protein-coupled receptor 48	to-to-hi-hi
	456716	AJ335440		glt3b2g11235142g10AA14470.1 Homo sapien	to-to-hi-hi
	407419			Target Exon	to-to-hi-hi
50	424363	AW512144	He.346947	ESTs, Weakly similar to A48809 carboxyle	to-to-hi-hi
	408966	AW292096	He.295036	glt3b2g11235142g10AA14470.1	to-to-hi-hi
	415516	F11411		glt3b2g11235142g10AA14470.1	to-to-hi-hi
	423144	AW851527	He.253677	ESTs, Weakly similar to I38022 hypothet	to-to-hi-hi
	452590	BE077084	He.99969	ESTs	to-to-hi-hi
55	433827	AA846538	He.187389	ESTs	to-to-hi-hi
	419709	AA255592	He.347973	ESTs, Weakly similar to alternatively sp	to-to-hi-hi
	413672	BE155536		glt3b2g11235142g10AA14470.1 Homo sapien	to-to-hi-hi
	452921	AA354572		glt3b2g11235142g10AA14470.1 Homo sapien	to-to-hi-hi
	427403	AA402107	He.257146	ESTs, Moderately similar to I38022 hypoth	to-to-hi-hi
60	433911	AW537461	He.255377	ESTs	to-to-hi-hi
	432583	AD427771	He.117170	ESTs	to-to-hi-hi
	448490	AIE73897	He.271692	ESTs, Weakly similar to I38022 hypothet	to-to-hi-hi
	448539	W80363	He.58446	ESTs	to-to-hi-hi
65	458382	AW778811	He.514451	ESTs, Weakly similar to ALU1_HUMAN ALU S	to-to-hi-hi
	453407	N02114		glt3b2g11235142g10AA14470.1 Homo sapien	to-to-hi-hi
	423231	AA323486	He.271273	Homo sapiens cDNA FLJ122336, tis, clone MA	to-to-hi-hi
	450628	AW382884	He.204715	ESTs	to-to-hi-hi
	411890	AA869253	He.136075	RNA, U2 small nuclear	to-to-hi-hi
	414730	U83567	He.77106	specfic, alpha, non-erythrocytic 1 (alp	to-to-hi-hi
70	444169	AV848170	He.587596	ESTs	to-to-hi-hi
	420911	U77413	He.100293	O-linked N-acetylglucosamine (GlcNAc) tr	to-to-hi-hi
	422196	AB007903	He.113062	KIAA0443 gene product	to-to-hi-hi
	452704	AA327823	He.149424	Homo sapiens PNAS-130 mRNA, complete cds	to-to-hi-hi
	425014	AA856100		Homo sapiens cDNA, FLJ122165, tis, clone H	to-to-hi-hi
75	426376	NA8752	He.302985	ESTs	to-to-hi-hi
	447754	AW073310	He.163533	Homo sapiens cDNA FLJ14142, tis, clone MA	to-to-hi-hi
	413856	AJ85213	He.71404	ESTs	to-to-hi-hi
	449000	U93650	He.38026	beta-tubulin protein CSPI1	to-to-hi-hi
	430064	AK000091	He.231436	hypothetical protein FLJ20384	to-to-hi-hi
80	412205	N33818	He.20274	ESTs, Weakly similar to unnamed protein	to-to-hi-hi
	423955	AJ205882	He.136164	cutaneous T-cell lymphoma-associated hum	to-to-hi-hi
	459519	BE063863		glt3b2g11235142g10AA14470.1 Homo sapien	to-to-hi-hi

	408722	AA487860	Hs.258102	ESTs	to-to-hi-hi
	459710	AJ01696	Hs.121592	ESTs	to-to-hi-hi
	417918	AA209205	Hs.163754	hypothetical protein FLJ12606	to-to-hi-hi
5	427964	AF759512	Hs.284163	NM_022095? Homo sapiens hypothetical C2H	to-to-hi-hi
	424367	AF069377	Hs.284163	ANKRD2N protein	to-to-hi-hi
	427220	AF069377	Hs.173993	RNA binding motif protein 6	to-to-hi-hi
	410451	BE065687	gb:RC3-BT0316-270400-016-110 BT0316 Homo	to-to-hi-hi	
	430713		NIA_005165? Homo sapiens nuclear factor 1	to-to-hi-hi	
10	407218	AA094373	Hs.28505	ubiquitin-conjugating enzyme E2H (homolo	to-to-hi-hi
	448012	HT1673	Hs.223626	ESTs	to-to-hi-hi
	419612	HA082627	Hs.110613	KIAA0421 protein	to-to-hi-hi
	455272	BE148152	gb:RC4-IT0231-041199-012-b04 HT0231 Homo	to-to-hi-hi	
	401839		NM_005177? Homo sapiens AT Paxis, H+ trans	to-to-hi-hi	
15	440422	AW452696	Hs.130760	myosin phosphatase, target subunit 2	to-to-hi-hi
	436819	AA731746	Hs.120232	ESTs	to-to-hi-hi
	413644	BE154910	Hs.276793	ESTs, Weakly similar to Z195_HUMAN ZINC	to-to-hi-hi
	413939	AL047051	Hs.199961	ESTs, Weakly similar to ALU7_HUMAN ALU S	to-to-hi-hi
	448186	BE022100	Hs.209406	ESTs, Weakly similar to I38500 zinc finger	to-to-hi-hi
	450468	AA009099	Hs.59169	ESTs, Moderately similar to HPV16 E1 pro	to-to-hi-hi
	433507	AI817336	Hs.191791	ESTs	to-to-hi-hi
	438996	AW748336	Hs.110613	KIAA0421 protein	to-to-hi-hi
	442769	AW904361	Hs.131191	ESTs, Weakly similar to ALU7_HUMAN ALU S	to-to-hi-hi
	427251	UE7611		transcendase 1	to-to-hi-hi
20	405051	AA080912		glucosylated 1 Stratiagene N1/T neuron (S67	to-to-hi-hi
	409123	AA063403		glucosylated 1 Stratiagene cerebellar stroma	to-to-hi-hi
	418225	AA577730	Hs.188684	ESTs, Weakly similar to P04259 ferritin	to-to-hi-hi
	433735	AA080965	Hs.109653	ESTs	to-to-hi-hi
	424404	AW443034	Hs.285578	ESTs	to-to-hi-hi
	446667	BE161878	Hs.224806	ESTs	to-to-hi-hi
	447982	H22953	Hs.137551	ESTs	to-to-hi-hi
	438990	AA527756	Hs.135049	ESTs, Weakly similar to ALU7_HUMAN ALU S	to-to-hi-hi
	427882	AA040987	Hs.193767	ESTs	to-to-hi-hi
	459580	H06982	Hs.423291	ESTs	to-to-hi-hi
25	416832	H09480	Hs.141304	ESTs	to-to-hi-hi
	453876	AW021748	Hs.110406	ESTs, Weakly similar to I38022 hypothetical	to-to-hi-hi
	414528	AI148650	Hs.188336	ESTs	to-to-hi-hi
	419802	AA084108	Hs.118920	ESTs	to-to-hi-hi
	409542	AA503020	Hs.36563	hypothetical protein FLJ22418	to-to-hi-hi
30	433500	AB25196	Hs.130891	hypothetical protein MG04400	to-to-hi-hi
	447499	AV026590	Hs.147674	prolactin receptor like 16	to-to-hi-hi
	435023	AB02552		gb:U7312.1 NCL_CGAP_Lc24 Homo sapiens	to-to-hi-hi
	412156	H24987	Hs.171110	Homo sapiens mRNA: cDNA DKFZp334C016 (f	to-to-hi-hi
	414505	AA5389	Hs.23558	ESTs, Weakly similar to AA0404 lysosomal	to-to-hi-hi
35	404277			NM_019111? Homo sapiens major histocompa	to-to-hi-hi
	414662	AL036658	Hs.78907	major histocompatibility complex, class	to-to-hi-hi
	444340	AI011153	Hs.6053	Homo sapiens cDNA: FLJ22783 fs, clone K	to-to-hi-hi
	445612	N94126	Hs.12889	hypothetical protein	to-to-hi-hi
	403739			ENSF00000251553? UDP-glucuronosyltransfe	to-to-hi-hi
40	402740			NM_001076? Homo sapiens UDP glycosylase	to-to-hi-hi
	411084	T18867	Hs.125472	ESTs, Moderately similar to KIAA0677 pro	to-to-hi-hi
	428143	AA333327	Hs.197335	plasma glutamate carboxypeptidase	to-to-hi-hi
	433060	D78874	Hs.8944	procollagen C-endopeptidase enhancer 2	to-to-hi-hi
	422749	W01078	Hs.278573	CD59 antigen p18-20 (antigen identified	to-to-hi-hi
45	428441	AJ224172	Hs.225096	lysozyme B (cathepsin family member)	to-to-hi-hi
	414382	AW303339	Hs.8065	hematopoietic P51-interacting protein	to-to-hi-hi
	441550	F13388	Hs.7888	Homo sapiens clone 23735 mRNA sequence	to-to-hi-hi
	446106	AA377165	Hs.44333	ESTs	to-to-hi-hi
50	482236	AI079378	Hs.170121	protein tyrosine phosphatase, receptor I	to-to-hi-hi
	446874	AW058304	Hs.58156	ESTs	to-to-hi-hi
	412795	BE241753	Hs.74592	special AT-rich sequence binding protein	to-to-hi-hi
	430325	AF004562	Hs.239356	synkin binding protein 1	to-to-hi-hi
	426392	AF096324	Hs.17384	ESTs	to-to-hi-hi
	447448	BE344265		F-box only protein 29	to-to-hi-hi
55	415743	AA167664	Hs.14333	ESTs, Weakly similar to Z195_HUMAN ZINC	to-to-hi-hi
	431607	AB033097	Hs.183669	KIAA1271 protein	to-to-hi-hi
	411918	X68134	Hs.72884	minkcat kinase-binding protein 5	to-to-hi-hi
	438320	BE36163	Hs.52025	ESTs, Weakly similar to ALU5_HUMAN ALU S	to-to-hi-hi
	431099	Y13367	Hs.249235	phosphotransferase 3, kinase, class 2, alpha	to-to-hi-hi
60	421687	AL035306	Hs.106823	hypothetical protein MG014797	to-to-hi-hi
	439565	AF068386	Hs.145599	ESTs	to-to-hi-hi
	442249	W00516	Hs.132355	Homo sapiens cDNA: FLJ22119 fs, clone H	to-to-hi-hi
	410096	AW245200	Hs.267400	hypothetical protein MG02540	to-to-hi-hi
	429447	AW812452	Hs.83286	ESTs, Weakly similar to S14747 sphingomy	to-to-hi-hi
65	431802	AL133570	Hs.270671	Homo sapiens mRNA: cDNA DKFZp334L201 (f	to-to-hi-hi
	411715	AB23453	Hs.242655	Homo sapiens cDNA: FLJ13289 fs, clone OV	to-to-hi-hi
	458230	BE311851	Hs.6636	KIAA1624 protein	to-to-hi-hi
	428788	AF022283	Hs.193516	B-actin CLL4 lymphoma 10	to-to-hi-hi
	450818	AI740573	Hs.142827	P311 protein	to-to-hi-hi
70	419578	AF002660	Hs.91251	hypothetical protein FLJ11198	to-to-hi-hi
	404001	AF159093		Homo sapiens endogenous retrovirus RAN1	to-to-hi-hi
	427004	AI021573	Hs.213107	ESTs	to-to-hi-hi
	401178	AA046772		RNA binding motif protein, X chromosome	to-to-hi-hi

WO 02/098358

PCT/US02/17594

	423749	U09848	Hs.132390	zinc finger protein 36 (K0X 18)	to-to-hi-to
	428558	AB033070	Hs.194408	KIAA1244 protein	to-to-hi-to
	435238	AW405546	Hs.127971	ESTs	to-to-hi-to
	428621	BE648798	Hs.55949	ESTs	to-to-hi-to
5	402155			Target Exon	to-to-hi-to
	415961	H10983	Hs.155919	ESTs	to-to-hi-to
	467265	AB023212	Hs.225967	KIAA0955 protein	to-to-hi-to
	412419	AW949830		gb:U6-F1001-05500-226-g05 F10001 Homo	to-to-hi-to
10	436397	AA095478	Hs.123206	ESTs	to-to-hi-to
	440509	BE410132	Hs.134202	ESTs, Weakly similar to T17279 hypoph	to-to-hi-to
	423895	AA332215		gb:EST136124 Embryo, 8 week 1 Homo sapien	to-to-hi-to
	402051			NM_006591 Homo sapiens ubiquitin spact	to-to-hi-to
	445504	AW296163	Hs.147296	ESTs	to-to-hi-to
15	423523	AK001409	Hs.274356	hypothetical protein FLJ10547	to-to-hi-to
	444290	AA262496		gb:zs2011.1 N1 CL_GCAP_GCB1 Homo sapien	to-to-hi-to
	433903	Z44194	Hs.4964	transducer of EREB2, 2	to-to-hi-to
	436995	N31273	Hs.42380	ESTs	to-to-hi-to
	401849			Target Exon	to-to-hi-to
20	402249			C19005533-g112741444ref(XP_008886.2)	to-to-hi-to
	406180	AB018249	Hs.170527	small inducible cytokine subfamily A (Cy	to-to-hi-to
	448176	AB72246	Hs.62184	ESTs	to-to-hi-to
	452629	AW938430	Hs.303303	hypothetical protein FLJ20618	to-to-hi-to
	457335	AW969834		ESTs	to-to-hi-to
	452444	BE144022		gb:MRQ-HT0165-191199-004-05 HT0165 Homo	to-to-hi-to
25	454249			Target Exon	to-to-hi-to
	433103	AA465299	Hs.124623	gb:sa23603.1 N1 CL_GCAP_GCB1 Homo sapien	to-to-hi-to
	439944	AA86767		ESTs	to-to-hi-to
	411283	AW852754		gb:PM1-CT0247-180190-009-03 CT0247 Homo	to-to-hi-to
	453195	R10085	Hs.130370	ESTs	to-to-hi-to
30	452554	BE004783		gb:MRQ-BN0114-270400-004-e11 BN0114 Homo	to-to-hi-to
	425884	AF020986	Hs.158201	Thymosin, beta 4, Y thymosone	to-to-hi-to
	429452	AY949495	Hs.133968	Homo sapiens cDNA FLJ133022, clone NT	to-to-hi-to
	431709	AF220185	Hs.267923	uncharacterized hypophalamus protein HT0	to-to-hi-to
	411701	BE181659		gb:DV1-HT0538-075506-191-g7 HT0538 Homo	to-to-hi-to
35	407029	AF722590	Hs.301283	KIAA0755 gene product	to-to-hi-to
	447476	BE293466	Hs.20880	ESTs, Weakly similar to I36022 hypoph	to-to-hi-to
	450436	AW293961	Hs.131887	ESTs	to-to-hi-to
	453535			CX001212-g77861632gb:AA70445.11 (AF2	to-to-hi-to
40	418635	AA244416		gb:zs2011.1 N1 CL_GCAP_P11 Homo sapien	to-to-hi-to
	446103	U09198	Hs.133904	hypothetical protein dM450223.2	to-to-hi-to
	400986			NM_024087 Homo sapiens hypothetical pro	to-to-hi-to
	424194	BE245833	Hs.169854	gb:TCBAP1E1908 Pediatric pro-B cell acut	to-to-hi-to
	402010			Exon Control	to-to-hi-to
45	402024			NM_005336 Homo sapiens high density lipo	to-to-hi-to
	400235			NM_005336 Homo sapiens high density lipo	to-to-hi-to
	405387			NM_022170 Homo sapiens Williams-Beuren	to-to-hi-to
	433075	NM_002959		sortilin 1	to-to-hi-to
	456332			C11000022-g77459103gb:JT20903 hypoph	to-to-hi-to
50	428181	AA423976		gb:zv62y06.e1 Soares_testis_NHT Homo sap	to-to-hi-to
	466629	AW891965	Hs.273789	histone deacetylase 3	to-to-hi-to
	426940	AA393537	Hs.98347	ESTs, Weakly similar to JC5308 testis-s	to-to-hi-to
	433565	AA539902	Hs.144211	Homo sapiens HERC2P7 pseudogene, partial	to-to-hi-to
	421431	AA650117	Hs.283107	ESTs	to-to-hi-to
	448631	AA54923		gb:hs3h12.x1 Soares_NFL_T_GDC_S1 Homo s	to-to-hi-to
55	433521	T69087	Hs.112482	Homo sapiens unknown mRNA sequence	to-to-hi-to
	407187	AA449971		gb:zv62y11.e1 Soares_testis_NHT Homo sap	to-to-hi-to
	498719	AT732707	Hs.116506	ESTs, Weakly similar to ALU7, HUMAN ALU S	to-to-hi-to
	440004	BE397117	Hs.120824	hypothetical protein FLJ21945	to-to-hi-to
	403947	NM_006032		piasin 3 (T isoform)	to-to-hi-to
60	405529	AW104558		chromosome 11 open reading frame2	to-to-hi-to
	402163			C11001075-g774557173gb:AA223607.1 HACO	to-to-hi-to
	404663			ENEP00000225664:KAA1521 protein (Fragme	to-to-hi-to
	400220			Exon Control	to-to-hi-to
65	401444			Target Exon	to-to-hi-to
	453824	BE143703		gb:MRQ-HT0164-191199-004-003 HT0164 Homo	to-to-hi-to
	400206			Exon Control	to-to-hi-to
	468559	AW748955	Hs.332520	Homo sapiens mRNA; cDNA DKFZp434A1014 (f	to-to-hi-to
	426966	AL080190	Hs.185242	Homo sapiens mRNA; cDNA DKFZp434A202 (fr	to-to-hi-to
70	428442	AA428635	Hs.98906	ESTs	to-to-hi-to
	440151	AA68167		gb:ak38e07.s1 Soares_testis_MIT Homo sap	to-to-hi-to
	431046	AW854382	Hs.249126	Homo sapiens clone 24894 mRNA sequence	to-to-hi-to
	443914	AI091173	Hs.222862	ESTs, Weakly similar to p40 [H.sapien]	to-to-hi-to
	402469			Target Exon	to-to-hi-to
75	418155	RA5481	Hs.23719	ESTs, Weakly similar to I35022 hypoph	to-to-hi-to
	446953	AI610818	Hs.7110	ESTs	to-to-hi-to
	423236	AW340958	Hs.7672	ESTs	to-to-hi-to
	421290	NM_014368	Hs.103137	L1M homeobox protein 6	to-to-hi-to
	450374	AA397540	Hs.90293	Homo sapiens clone 122482 unknown mRNA	to-to-hi-to
	402247			Target Exon	to-to-hi-to
80	415184	AA380436	Hs.211973	homolog of Yeast FRP4 (ribosomal RNA pro	to-to-hi-to
	415632	U67085	Hs.78324	Tot37 homolog	to-to-hi-to
	423718	AL119520	Hs.180737	Homo sapiens clone 23664 and 23905 mRNA	to-to-hi-to









WO 02/098358

PCT/US02/17594

	414483	R25513	Hs.10683	ESTs	hi-hi-hi
	451273	NM_014811	Hs.26163	KIAA0549 gene product	hi-hi-hi
	437052	AA861697	Hs.120591	ESTs	hi-hi-hi
	443049	R06959	Hs.15769	hypothetical protein MGC4174	hi-hi-hi
5	429483	A0874832	Hs.128708	ESTs	hi-hi-hi
	411296	BE207307	Hs.10114	growth suppressor 1	hi-hi-hi
	425188	AK002052	Hs.155071	hypothetical protein FLJ11190	hi-hi-hi
	435315	BE360513	Hs.27535	hypothetical protein MGC4837	hi-hi-hi
	402327	AI127076	Hs.306201	hypothetical protein DKF2p564C1278	hi-hi-hi
10	431089	BE041395	Hs.53542	ESTs, Weakly similar to unknown protein	hi-hi-hi
	418824	AW751661	Hs.53542	choreoacanthocytosis gene; KIAA0908 prot	hi-hi-hi
	443926	AB002365	Hs.23311	KIAA0367 protein	hi-hi-hi
	450149	AW69781	Hs.120933	Zn family member 2 (pkl-paired Oncoferrin)	hi-hi-hi
	418443	NM_005239	Hs.85146	v-avl avian erythroblastosis virus E26 o	hi-hi-hi
15	468992	BE549005	Hs.231754	ESTs	hi-hi-hi
	410102	AW248508	Hs.279727	ESTs, homologue of PEM3 (Clona scapnyi)	hi-hi-hi
	451062	AL110125	Hs.25910	Homo sapiens mRNA; cDNA DHFZp594C1416 (f	hi-hi-hi
	407633	NM_007039	Hs.57199	similar to rat HFEV107	hi-hi-hi
20	418941	AA452970	Hs.239527	E1B-55kDa-associated protein 5	hi-hi-hi
	407059	X95408	-	gcb-Laspeiro cyclin E gene.	hi-hi-hi
	455956	BE162704	-	gcbPM1-HIT0454-30 c299-001-c28 HIT0454 Homo	hi-hi-hi
	437783	AA463969	Hs.5331	tissue inhibitor of metalloproteinase 1	hi-hi-hi
	451404	AA460775	Hs.6295	ESTs, Weakly similar to T17248 hypotheti	hi-hi-hi
25	428494	AA233439	Hs.184634	hypothetical protein FLJ20005	hi-hi-hi
	414367	DE1283	Hs.45206	ESTs	hi-hi-hi
	458415	A1734051	Hs.277102	ESTs, Weakly similar to ALU1_HUMAN ALU S	hi-hi-hi
	400183	-	-	Eos Control	hi-hi-hi
	400158	-	-	ENSP00000240320-cDNA FLJ11591.1; clone	hi-hi-hi
	403893	-	-	ENSP0000023706P-Protocadherin alpha 6 p	hi-hi-hi
30	423869	A1223833	Hs.154463	ESTs	hi-hi-hi
	400170	-	-	Eos Control	hi-hi-hi
	403291	-	-	Target Exon	hi-hi-hi
	422026	U80736	Hs.110826	intructeotide repeat containing 9	hi-hi-hi
35	417130	AW276558	Hs.81226	S100 calcium-binding protein A4 (calcium	hi-hi-hi
	423472	AA546781	Hs.136418	ESTs	hi-hi-hi
	405231	-	-	C2001066g(1025742)g(MP_033892.1) CD	hi-hi-hi
	400141	-	-	Eos Control	hi-hi-hi
	429971	BE276404	Hs.258113	hypothetical protein FLJ11887	hi-hi-hi
40	422360	AW450583	Hs.124130	ESTs, Weakly similar to T23682 hypotheti	hi-hi-hi
	425538	BE270918	Hs.164028	Homo sapiens, clone IMAGE3534875, mRNA,	hi-hi-hi
	455972	A054347	Hs.2017	ribosomal protein L38	hi-hi-hi
	456622	AF205649	Hs.107740	Kruppel-like factor 2 (lung)	hi-hi-hi
	418615	A058453	Hs.19487	hypothetical protein similar to CHH_HUMAN CORN1	hi-hi-hi
45	448439	BE113082	Hs.28229	APC29 protein	hi-hi-hi
	445418	AW139377	Hs.127179	cryptic gene	hi-hi-hi
	402559	Z23024	-	Rho GTPase activating protein 1	hi-hi-hi
	402575	Z23024	-	Rho GTPase activating protein 1	hi-hi-hi
	420811	AA075544	-	ESTs, Weakly similar to S34323 GTP-bind	hi-hi-hi
50	446827	A973016	Hs.15725	hypothetical protein SBH48	hi-hi-hi
	400247	-	-	Eos Control	hi-hi-hi
	430289	AK001952	Hs.238039	hypothetical protein FLJ11090	hi-hi-hi
	400133	-	-	Eos Control	hi-hi-hi
	418816	T29621	Hs.88778	carboxyl reductase 1	hi-hi-hi
55	433579	BE264473	Hs.284297	hypothetical protein from EUROIMAGE 1967	hi-hi-hi
	401952	-	-	Target Exon	hi-hi-hi
	410349	AW663021	Hs.323445	ESTs, Weakly similar to T2D3_HUMAN TRANS	hi-hi-hi
	417559	AF045229	Hs.82280	regulator of G-protein signalling 10	hi-hi-hi
	446851	AW007332	Hs.10450	Homo sapiens cDNA: FLJ22063; clone H	hi-hi-hi
60	404489	-	-	Target Exon	hi-hi-hi
	405802	-	-	Target Exon	hi-hi-hi
	455266	L26073	Hs.198726	cold shock domain protein A	hi-hi-hi
	457133	ME4968	-	v-Ki-ras2 Kirsten rat sarcoma 2 viral on	hi-hi-hi
	458330	C16931	-	gbcC16931 Clontech human aorta polyA mRNA	hi-hi-hi
65	433041	BE265948	Hs.289030	colon cancer-associated protein Mcl1	hi-hi-hi
	445645	AK11786	Hs.164192	ESTs, Weakly similar to Y157_HUMAN HYPOT	hi-hi-hi
	414911	NM_000107	Hs.77602	damage-specific DNA binding protein 2 (f	hi-hi-hi
	414682	AL021154	Hs.76884	inhibitor of DNA binding 3, dominant neg	hi-hi-hi
	422311	A1073815	Hs.114948	cytokine receptor like factor 1	hi-hi-hi
	447329	BE390517	-	ESTs, Moderately similar to ALU1_HUMAN A	hi-hi-hi
70	412842	AL120344	Hs.75074	mitogen-activated protein kinase-activat	hi-hi-hi
	420747	BE294407	Hs.99910	phosphotransferase, platelet	hi-hi-hi
	431912	AB60552	Hs.76549	ESTs, Weakly similar to A56154 AM subo	hi-hi-hi
	446506	A1021118	Hs.15159	chemokine-Rho factor, alternatively sp	hi-hi-hi
	408933	AW963172	-	PRC2000 protein	hi-hi-hi
75	433875	AW977653	Hs.75319	ribonucleotide reductase M2 polypeptide	hi-hi-hi
	424560	AA158727	Hs.150505	protein predicted by clone Z3735	hi-hi-hi
	425234	AW152225	Hs.165909	ESTs, Weakly similar to C8022 hypotheti	hi-hi-hi
	439615	A3206079	Hs.66693	hypothetical protein FLJ20420	hi-hi-hi
	410174	AA306007	Hs.59461	DKFZP434C245 protein	hi-hi-hi
80	410442	X73424	Hs.63788	propionyl Coenzyme A carboxylase, beta p	hi-hi-hi
	429190	H18650	Hs.92862	ESTs	hi-hi-hi
	423619	T49891	Hs.249159	adrenergic, alpha-2v-, receptor	hi-hi-hi

WO 02/098358

PCT/US02/17594

433794	AW753679	Ha.30962	ESTs	hi-to-hi
421998	R74441	Ha.117175	poly(A)-binding protein, nuclear 1	hi-to-hi
451593	AF151879	Ha.28708	CG-121 protein	hi-to-hi
452292	BE246374	Ha.27842	hypothetical protein FLJ11210	hi-to-hi
447425	AB57147	Ha.1874578	acylphosphatase 1, cytosolic (common)	hi-to-hi
421554	AW153267	Ha.106459	suppressor of var1 (S.cerevisiae) 3-like	hi-to-hi
432502	NM_014641	Ha.277585	KGA0170 gene product	hi-to-hi
429597	NM_003818	Ha.2442	a disintegrin and metalloproteinase domain	hi-to-hi
434203	BE252877	Ha.283558	hypothetical protein PRO1055	hi-to-hi
438491	AW074465	Ha.260349	phosphoserine aminotransferase	hi-to-hi
409142	AL138877	Ha.50758	SMC4 (structural maintenance of chromosome	hi-to-hi
436774	A0498788	Ha.165190	ESTs	hi-to-hi
438182	AW342140	Ha.182545	ESTs, Weakly similar to ALU1_HUMAN ALU S	hi-to-hi
449103	T24968	Ha.23270	HSPC71 protein	hi-to-hi
421059	AB54133	Ha.30212	thyroid receptor interacting protein 15	hi-to-hi
449939	AL133363	Ha.16906	CG-32 protein	hi-to-hi
408576	NM_003542	Ha.45423	H4 histone family, member G	hi-to-hi
410073	AW408163	Ha.58458	calsitin (calyculin-associated protein), a	hi-to-hi
450912	AW530251	Ha.25547	v-hes FBJ murine osteosarcoma viral onco	hi-to-hi
434701	AA450479	Ha.321707	KGA0742 protein	hi-to-hi
450465	AL117424	Ha.20335	chloride intracellular channel 4	hi-to-hi
451144	AW595103	Ha.61712	pyruvate dehydrogenase kinase, isoenzyme	hi-to-hi
427380	AK521153	Ha.289231	Homo sapiens cDNA: FLJ21111 fls, clone L	hi-to-hi
451831	NM_001074	Ha.480	activating transcription factor 3	hi-to-hi
406776	T16208	Ha.237164	ESTs, Highly similar to LDH1_HUMAN L-LAC	hi-to-hi
428157	A0758719	Ha.194427	hexokinase 2	hi-to-hi
409396	BE259182	Ha.82765	glyoxylate reductase	hi-to-hi
416203	X54942	Ha.83758	CDC28 protein kinase 2	hi-to-hi
449338	H73444	Ha.394	adrenomedullin	hi-to-hi
422382	AA016188	Ha.111244	hypothetical protein	hi-to-hi
429707	A073225	Ha.41270	prothymosin-yalin, 2-oxoglutarate 5-4-0	hi-to-hi
416555	AW598813	Ha.79428	BCL2/adrenomedullin E1B 19kD-interacting pro	hi-to-hi
419551	AW582266	Ha.91011	anterior gradient 2 (Xenopus laevis) hom	hi-to-hi
434094	AA305699	Ha.238226	hypothetical protein PRO2013	hi-to-hi
443951	F13272	Ha.111334	ferritin, light polypeptide	hi-to-hi
422975	AA347720	Ha.122839	KGA0254 protein	hi-to-hi
403314	AA369601	Ha.239138	pro-B cell colony-enhancing factor	hi-to-hi
412664	AA421404	Ha.346898	nucleolar protein p40; homolog of yeast	hi-to-hi
409089	T09799	Ha.42544	thioridone-like	hi-to-hi
409660	V45393	Ha.55568	activator transcription factor 7	hi-to-hi
424332	A1893251	Ha.8248	Target CAT	hi-to-hi
408388	AF091086	Ha.44553	hypothetical protein	hi-to-hi
441252	AW360901	Ha.183047	hypothetical protein MGC4399	hi-to-hi
433369	X7732	Ha.51184	nucleolinin 2	hi-to-hi
443837	A1984025	Ha.9864	spindle pole body protein	hi-to-hi
426108	AA622037	Ha.166468	programmed cell death 5	hi-to-hi
441181	AA418925	Ha.121076	peptidylprolyl isomerase (cyclophilin)-4	hi-to-hi
447397	BE247876	Ha.18442	E1 enzyme	hi-to-hi
427505	AA361592	Ha.178781	26S proteasome-associated pad1 homolog	hi-to-hi
430287	AW182459	Ha.125769	ESTs, Weakly similar to LEU5_HUMAN LEUKE	hi-to-hi
415857	AA856115	Ha.127797	Homo sapiens cDNA FLJ11381 fls, clone HE	hi-to-hi
423198	X81933	Ha.1634	cell division cycle 25A	hi-to-hi
427867	A0329211	Ha.375558	hypothetical protein FLJ11149	hi-to-hi
431374	BE258632	Ha.251871	CTP synthase	hi-to-hi
413273	U75678	Ha.75257	stem-loop (histone) binding protein	hi-to-hi
442739	AA54739	Ha.88505	ESTs	hi-to-hi
443881	RA5412	Ha.237146	hypothetical protein FLJ12762	hi-to-hi
416209	AA236776	Ha.75078	MAD2 (mitotic arrest deficient, yeast, h	hi-to-hi
421834	BE543205	Ha.288771	DKFZP583A0222 protein	hi-to-hi
411263	BE287802	Ha.69360	kinesin-like 6 (mitotic centromere-asso	hi-to-hi
413824	AL119854	Ha.75916	seisun-1	hi-to-hi
463698	AF151078	Ha.261199	hypothetical protein	hi-to-hi
438453	BE254874	Ha.55566	thyroid hormone receptor interactor 13	hi-to-hi
429912	A062649	Ha.262587	pituitary tumor-transforming 1	hi-to-hi
443425	A0768158	Ha.9329	chromosome 20 open reading frame 1	hi-to-hi
422563	C16165	Ha.201191	epithelial membrane protein 2	hi-to-hi
419879	Z17805	Ha.93354	Homeo, neuronal immediate early gene, 2	hi-to-hi
422383	T55979	Ha.115474	replication factor C (activator 1) 3 (38	hi-to-hi
416065	BE287931	Ha.78998	proliferating cell nuclear antigen	hi-to-hi
424356	AW175831	Ha.154443	mitochondrion maintenance deficient (S.	hi-to-hi
447519	U62558	Ha.336955	ESTs	hi-to-hi
437979	NM_014214	Ha.5753	inositol(myo)-1(or 4)-monophosphatase 2	hi-to-hi
443636	AC020263	Ha.15797	ciron (rho-interacting, sesin/thrombin	hi-to-hi
422044	AF129535	Ha.272327	F-box only protein 5	hi-to-hi
440324	BE275112	Ha.71156	zinc finger protein 259	hi-to-hi
421921	H83363	Ha.5820	translocase of inner mitochondrial mem	hi-to-hi
422638	NM_0010809	Ha.1594	centromere protein A (17kD)	hi-to-hi
427719	A0391322	Ha.134728	ESTs	hi-to-hi
422263	AA411507	Ha.114311	CDC45 (cell division cycle 45, S. cerevis	hi-to-hi
424840	D79897	Ha.134719	extra spindle poles, S. cerevisiae, homo	hi-to-hi
418216	AA682240	Ha.283099	AF15c14 protein	hi-to-hi
412140	AA219891	Ha.73525	RAB8 interacting, kinesin-like (rakibios	hi-to-hi



WO 02/098358

PCT/US02/17594

	412722	AI343300	Hs.15091	ESTs	hi-to-hi
	446839	BE091526	Hs.16244	mitotic spindle coiled-coil related prot	hi-to-hi
	428962	NM_000346	Hs.2316	SRY (sex determining region Y)-box 9 (ca	hi-to-hi
	435108	AI163534	Hs.6467	synaptophysin 3	hi-to-hi
5	430178	AIW44812	Hs.152475	ESTs	hi-to-hi
	421733	AL119671	Hs.1420	fibroblast growth factor receptor 3 (ach	hi-to-hi
	462410	AL133619		Homo sapiens mRNA; cDNA DKFp434E2321 (f	hi-to-hi
	430132	AA204696	Hs.234149	hypothetical protein FLJ20647	hi-to-hi
	428297	AA236251	Hs.193593	zeste (or cyclin) proteinase inhibitor	hi-to-hi
10	413142	M81740	Hs.75212	omithine decarboxylase 1	hi-to-hi
	427239	BE270447	Hs.174070	ubiquitin carrier protein	hi-to-hi
	406738	BE222875	Hs.56205	insulin induced gene 1	hi-to-hi
	410748	BE353816	Hs.12532	chromosome 1 open reading frame 21	hi-to-hi
15	424506	AF220490	Hs.149623	group III acorated phospholipase A2	hi-to-hi
	447333	BE090580	Hs.70704	hypothetical protein dJ61668.3	hi-to-hi
	414761	ALU77228	Hs.77256	enhancer of zeste (Drosophila) homolog 2	hi-to-hi
	415902	AIW24854	Hs.91521	hypothetical protein	hi-to-hi
	411665	BS112676	Hs.303116	stromal cell-derived factor 2-like 1	hi-to-hi
20	482322	BE566343	Hs.28983	glutaredoxin (thioltransferase)	hi-to-hi
	426006	RG0031	Hs.22627	ESTs	hi-to-hi
	457465	AIW301344	Hs.122908	DNA replication factor	hi-to-hi
	404867	AA157857	Hs.182265	keratin 19	hi-to-hi
	407230	AA157857	Hs.182265	keratin 19	hi-to-hi
25	446881	AJ003624	Hs.15896	keratin	hi-to-hi
	408453	BE236854	Hs.46039	phosphoglycerate mutase 2 (muscle)	hi-to-hi
	435166	AIW87274	Hs.105435	GDP-mannose 4,6-ephydratase	hi-to-hi
	424544	M88700	Hs.150403	dopa decarboxylase (aromatic L-amino aci	hi-to-hi
	431325	AIW026751	Hs.5794	ESTs, Weakly similar to 2109260A 3 cell	hi-to-hi
	414922	D00723	Hs.77631	glycine cleavage system protein H (mito	hi-to-hi
30	432821	BS214605	Hs.289952	Homo sapiens cDNA: FLJ2390 fs, clone H	hi-to-hi
	418574	N02754		U-phase phosphoprotein S	hi-to-hi
	409342	AIU07058	Hs.54089	BRCA1 associated RING domain 1	hi-to-hi
	432734	AA837306	Hs.263625	US1-interacting protein NUDE1, rat homo	hi-to-hi
	438087	BE300296	Hs.50594	CSK-133 protein	hi-to-hi
35	420309	AIW43637	Hs.21786	ESTs, Weakly similar to ALU5_HUMAN ALU S	hi-to-hi
	411619	AI119609	Hs.71040	hypothetical protein FLJ20425	hi-to-hi
	424381	AA285249	Hs.146329	protein kinase Chk2	hi-to-hi
	442547	AA306997	Hs.217484	ESTs, Weakly similar to ALU1_HUMAN ALU S	hi-to-hi
40	430376	AIW250253	Hs.12532	chromosome 1 open reading frame 21	hi-to-hi
	434866	AF151103	Hs.112259	T cell receptor gamma locus	hi-to-hi
	412330	NM_005100	Hs.738	A kinase (PRKA) anchor protein (gravin)	hi-to-hi
	462123	AI267615	Hs.39022	ESTs	hi-to-hi
	424863	AIW295112	Hs.139649	Homo sapiens cDNA FLJ13303 fs, clone OV	hi-to-hi
45	428907	AJ043641	Hs.185758	ESTs	hi-to-hi
	431566	AF176012	Hs.260720	J domain containing protein 1	hi-to-hi
	439979	AIW600291	Hs.6323	hypothetical protein FLJ10430	hi-to-hi
	418836	AI655459	Hs.161712	ESTs	hi-to-hi
	433767	AIW49974	Hs.152070	ESTs	hi-to-hi
50	425236	AIW067800	Hs.155223	stannocalcin 2	hi-to-hi
	426215	AIW963419	Hs.156223	stannocalcin 2	hi-to-hi

WO 02/098358

PCT/US02/17594

TABLE 2B

Play: Unique Eos probe/identifer number  
CAT number: Gene cluster number  
Accession: Genbank accession numbers

Play	CAT Number	Accession
10	403600 10729A-1	AA523775 AA056342 AA538078 AW975281 AA664986
	403601 109699-1	AA083912 AA075319 AA083403 AA076264 AA078992 AA084026 AA081361 AA113913 AA113882 AA083621 AA134801 AA082953 AA070343
	409123 110143-1	AA063403 AA070823 AA070060
	410216 119664-1	BE061839 AA059663 AA060665
	410461 120411A-1	BE065867 BE065537 AW742002 HT3690
	410496 120611-1	AA355749 AA085620 AW066333 AA340319 BE170036
15	411053 123046-1	AW181501 HT1965 AW181502 AW181504 AW181501 AW181504 BE152631 BE152490 BE149043 BE149075 BE149035 BE149087
	411233 123638-1	AW183373 AW183379 AW183346 AW183371 AW183375 AW183362 AW183367 AW183337
	411283 123769-1	AW182754 AW182807 AW182807 AW182817 AW182817 BE172795 AW183544
	411701 125446-1	BE181659 AW080676 AW187638
20	411831 126040-1	AW1994394 AW199439 AW199439 AW199439 AW199439 AW199439 AW199439 AW199439 AW199439 AW199439
	412419 126418-1	AW194830 AW194836 AW194836 AW194836 AW194836 AW194836 AW194836 AW194836 AW194836 AW194836
	412492 130082-1	AW194836 AW194836 AW194836 AW194836 AW194836 AW194836 AW194836 AW194836 AW194836 AW194836
	412657 1318507-1	AW197616 C04000
25	413361 136360-1	BE086815 BE086823 BE1218 BE1229
	413309 137433-1	BE145419 BE145433
	413672 1382512-1	DE166306 BE166439 BE166700 BE166449 BE166653 BE166533 BE166524 BE166670 BE166721 BE166723
	413308 1533673-1	F05251 N13748 Z4028 H14747
	415515 1539185-1	F11411 N15237 Z4381 H20760
30	416508 1597894-1	U39769 N13143 H60012
	416631 1605019-1	H94645 H9384 H96694
	416964 163427-1	A1222358 N13590 D51948 AA243520 AA190953
	417314 1666649-1	N68168 N69186 N60450
	418056 171841-1	AA524896 AW971347 AA211537
	418259 173358-1	AA215404 AW090899 BE164152 AW1271459 N74332 A126261
35	418574 176601-1	U28754 N16247 A168145 AW167769 A1232671 AW1955043 A1900326 A1776405 A1016250 AA843678 AW161882 N23137 N23129
		W70051 A1038748 A1831327 A1925845 AW458986
		AA244416 AA244401
		AA807544 AA200648 A1243066 AK22744 AA170289 AA829425 AW452096 AW829317 F19039 AA252024
		AA011741 AA300096
		AA3011270 AA301379 AA301366
40		AW181145 AA490718 H68637 AA304575 T06067 AA331991
		H09446 AA332597 AW594790 BE143680
		ALJ33533 F11784 H185104 T06069 H23079 F19495 AW134660 A129437 AL133896 AA057405 N78367 AW17450 A102852 T09262
		T09008 H23080 A1020874 AA894145 A1732387 A1791768 N1733447 AA688785 N62123 T09261 AW366835
45	423895 233006-1	AA332215 AA403110 AW966299
	424593 241234-1	AA343729 AA455779 AA343470
	425074 245498-1	AA196830 AW10890 H67851 AA350358 BE168712
	425291 249181-1	AA345472 AW052381 AW181319 AW161041 A1744649
	425980 257773-1	AA366951 AA470999 AA469425
50	426413 266650-1	AA377823 AW954494 A1022388
	428181 287953-1	AA422875 AA437075 BE050569
	429163 300542-1	AA584786 AW197421 AA523975 AA473512
	429640 305328-1	M65776 AA454536 AA456208 H10189
	430088 312849-1	AA454984 M85405 AA947568
	430103 313089-1	AA466259 AW597142 AW1897144
	430439 31808-1	AL133551 AL041080 AL117481 AL112260 AW430292 A1988026
	431089 321325-1	BE041395 AA491826 AA521946 AA176930 AA668102
	431843 338324-1	AA516420 A14818 C14815 C15161 F15088 D08076 D08066 AW970134 AA450007 D81004 D06184 A1498371 D03822 D0181 C15876
	432079 341114-1	AW972746 AA525232 A1503014
60	432340 345248-1	AA534222 AA532532 T01234
	432676 362892-2	A187366 AA536809 A1614878
		N10_02059 X98248 AA232378 AA164376 NA470560 AA170533 BE327147 AW19271 AA017125 A1199417 A1356213 A1168442 A1337018
		AA175049 A185459 AA195895 AA588000 AA418326 AA418378 N71981 AL043634 AA426361 AA418275 AA232975 A103861 BE277220 BE387505
		N89710 AW1675004 AA16228 AL079511 H85743 AW090219 AW090907 AA194368 T38210 AA054525 AA21752 AA058441 AW300669
		AW196218 T02287 BE464033 AW173548 A1890502 BE552385 A1890196 AW181831 A1238559 AW195065 AA0118359 A27377 AL042958
		AA411308 AA402810 H38111 AW013631 AW206432 AW752435 AW376124 A292520 A1292121 AA104047 BE613672 BE40874 BE545915
		BE517026 BE19688 AW402892 AW247466 BE52333 AA134761 BE54019 BE26105 BE313080 BE547713 BE538678 BE546749
		AA34184 BE17138 BE253377 BE7598 H29072 AA350680 BE076529 BE253657 AA532613 BE252486 AW084459 D03586 R8769 AW461832
		BE095388 AW086222 A1259455
		R76593 AF147360 R76594
		A182552 A135343 A1800510 A1577711 F424263 AA661876
70	433540 4272894-1	AA335450 AA27894 AA90581
	433616 457544-1	A182180 BE1106 AT44264 AA080846 A643417 AA543616 Z70715
	435882 42814-2	BE143433 AA071273 AA074787 AW187308 BE312102 AW149824 BE17985 AW197283 BE071945 BE072005 AW57355 BE071965 AW328321
	437516 43892-1	BE072000 BE071960 AW573306 AW174930 AW187302 X97303 AW099522 BE060019 BE062219 BE266665 BE264970
		AF075009 R53109 R53308
	438882 466469-1	AA527605 AA533754 AW187846
	439880 467544-1	AW182334 A1882567 A1803822
	439046 468133-1	AA947354 AA529660 A1857296
	439848 477806-1	AW1979249 D63277 A1846868
	440151 487109-1	AA868167 F21558 F31418 F35624
80	440507 496577-1	H09994 BE147598

WO 02/098358

PCT/US02/17594

5	441102 520604_1	AA973605 AI299888 AA017019 H63226 T90771
	442048 531432_1	AA974803 AI984319 AW304095
	443161 561305_1	A1038316 AI244651 AI261653
	444290 59964_1	AA262498 AW48929 AA303636 D61644 D78724
	444314 600667_1	AI140497 AW749625 AW749635 AW749644
10	445808 65133_1	AW652254 AW965332 AA304236
	447329 71759_1	BE090517 AW970792 AW064490 AW014985 F27436 AA94736F 15843 H89336 AA563626 F17712 BE566579 AA121821 AA284852 AA77751
	447448 722246_1	AI025246
	448150 752165_1	BE244286 CH6429 H42373 AI820705 AI379786 R55439 AW276142
	448489 765247_1	AI472167 AW903013 R32175
15	448631 772996_1	AI523876 RA5782 RA5781
	448736 77790_1	AI549823 AI02256
	452410 9155_1	BE014081 W01986 AW500790
		AI113615 AA481118 AA383604 AA76447 T09430 AI673758 AA524896 AI581545 AI300820 AW498812 AA256162 AI589724 AI685732 AA602400
		AA054545 AI204595 AW165541 AA157456 AA152629 AA383652 AA431072 AW582707 AA35410 AA727464 AI215594 AA622747 R74039
20	452444 918078_1	N35031 AI804128 AW513621 AA686361 AI026626 AA493386 AA614641 W81604 AI657080 AI214351 AA730140 AI212574 AI200813 AI269603
	462654 925931_1	AI562082 AI807055 AA765529 AA569809 AI368449 AI888077 AB562930 AW080338 AA575863 AA730154 AI767072 AA488516 AI734130 AI734138
	463754 925931_1	AA426394 AA432097 AT41241 AW043563 AI32741 AI732734 AA437305 AA423820 AA506408 R04130
	464264 925931_1	BE144022 BE143969 BE143915
	464764 925931_1	BE044783 BE040497 AI811790
25	465715 1234106_1	BE160229 AW196719 AW820179 AW819882 AW819876 AW820169 BE153201 AW959376 BE162911
	465719 1249136_1	AW850816 AW932633 AW851100
	465722 1271871_1	BE144152 BE145133 BE148159 BE148132 AW885107
	465619 1343307_1	BE063953 BE063955 BE063866 BE063705 BE063646 BE061416 BE063844
	465663 1348742_1	BE154078 BE153973 BE064801 BE163862 BE153847 BE064694 BE163602 BE068075 BE154078 BE064772 BE064942 BE163557 BE153509
30	465729 1353792_1	BE072032 BE072106 BE072056 BE072098 BE072103
	465824 1372880_1	BE143702 BE143631 BE143629 BE143702
	465956 1387163_1	BE162704 BE162705 BE162732 BE162702 BE162694
	466123 1534442_1	R00602 ZA2921 F06132
	467133 29066_1	NS4968 NM_004885 AB093924 AL135130 AW242010 AA470948 AT740449 MT7007 B03210 M35506 M35504 L00049 AI185396 W35273 X071669
35		X02258 W63635 AW56020 AI53945 AA452563 AA468811 W21091 T28576 AW677922 BE350180 AW684073 AI144938 AW117295 AA311229
		AI343010 AA766141 BE219368 AW6249 AA280396 AW905474 AA323870 AW770018 AA262948 AW450230 AW382390 AW690447 AW581613 AW699941
		AA425937 AW380065 AA330647 AA282180 T27356 H85307 AA861543 AA335548 AA356410 AW850556 AW380647 AW581613 AW690649
		AI657016 W10374 AW474707 AA605084 AA682136 AW949315 AA361728 N33863 AA411821 AA401640 AW949446 AI120766 AI600024
		AW771691 H44657 DE1651 AI338460 R14164 AI301629 N64676 AV569559 AI697660 AI064579 AA267927 AW453362 AW601642 AA767681
40		AA737710 AW572481 AA326104 AA542415 BE464969 BE045206 AW167917 AA433916 AA523501 AI015987 W25230 AI889491 AW773486
		AA937541 AI334416 AA76214 A281159 AA536309 AA582188 AA256527 AW166015 AA670007 H08199 AA030271 AA261015 W47527 AA468252
		AI364302 AA689265 R40473 H02312 AA648116 AA342730 AA243624 R99351 R41589 RA9696 AA854442 F01713 AA213686 AA721296 R79833
		H16241 R77088 H65554 AA220726 N65349 AI374613 AI306683 AA156089 AA195448 AA635570 AA72321 AI562715 AA195753 AA174563
		AW673045 AI229133 AI021612 AW74600 AW557807 AA462223 AA835804 T91295 H85138 AA352936 AW119152 N16774 AB85584 N39418
45		AA684877 AA679469 BE350951 N1020 AI060915 F00075 AA648786 N26790 AA828898 AW019991 AW796631 AW983262 N48632 AA564482
		AV654033 H54481 AW945712 C03289 AV655314 AV653070 AV659806 AV660435 H70113 C03233 R31984 196949 AV655936 AV658679
		H69137 AA384411 AA12534 C03249 W32014 R87686 C09526 BE33077 N24364 AA287961 N80109 F08452 R12740 H06297 AI138364
		AW020001 BE178443 BE178616 BE178336 BE178660 BE178107 BE178395 BE178215 BE178168 BE178447 BE178352 BE178422 BE178424
		BE178043 BE178093 BE178460 BE178356 BE178441 BE178438 BE178467 AW01259 BE177839 BE178094 R28455 BE177844 BE178100
50		AA262387 R20669 W0334 W36659 AA258711 BE178141 BE177893 BE178449 AA167718 H89694 BE178017 BE178029 BE177996 BE177936
		AA095144 N32462 AA281203 AA281183 W47526 W06015 R34165 R35396 T97366 R79640 W25259 R59450 AW368425 BE178196 R26447
		C03146 C03683
		U25750 AP702472 AA437376 AB872282 AA467262 R22383 AI865790 R21532 AA563628 AW571889 AA377111 R78814 T27193
		BE220675 AA45821 AA009962

TABLE 2C

Key: Unique number corresponding to an Ensembl probe

Ref: Sequence source. 7 digit numbers in this column are Genbank Identifier (GI) numbers. "Dunham I. et al." refers to the publication entitled "The DNA sequence of human chromosome 22" Dunham I. et al. (1999) Nature 402:489-495.

Strand: Indicates DNA strand from which exons were predicted.

NL\_position: Indicates nucleotide positions of predicted exons.

Key	Ref	Strand	NL_position
400481	8438853	Plus	112433-112541
400501	9796227	Minus	12479-12619
400713	8118874	Minus	43185-43394
400769	8113128	Plus	28571-28795
400818	8565994	Plus	172844-172765,173085-173200
400881	2842777	Minus	91445-91603,92123-92265
400892	2842777	Minus	110431-110709
400905	7770576	Minus	173043-173864
400988	8036487	Minus	63140-63319
400995	6099094	Plus	141166-141601
401093	8516137	Minus	22335-23166
401178	8438616	Minus	133653-133812
401192	8718952	Minus	69569-70101
401205	7712287	Plus	164832-165112
401405	7766126	Minus	69276-69452,69548-69968
401416	7452659	Minus	121456-121626
401419	7452659	Plus	130369-130606
401444	8346725	Plus	90956-90994,93070-93213
401512	7622346	Plus	136399-136557
401563	6247910	Plus	91356-91763
401600	6387146	Minus	27363-27516,28727-28851,29636-29731
401750	9228551	Plus	67143-67270,85284-88373,90596-90770,95822-96001,96688-96775,96870-96992,98048-96136
401757	7236630	Plus	66641-88761
401836	7656637	Plus	1016-1068,2751-2667,3241-3346,26677-26631
401846	7770425	Plus	125375-129483,129567-129720
401852	3315121	Minus	63710-63875
401868	3126761	Plus	22397-29916
402082	8117478	Minus	180046-180183
402101	8117897	Plus	134306-134467,136462-136667,136421-136548
402108	6191621	Plus	37117-3846
402163	8568936	Plus	165596-167119
402165	6576002	Plus	25466-25639
402240	7660131	Plus	104362-104527,106136-106372
402249	7704953	Plus	107636-107813,10894-10924,110435-110502,113182-113886
402347	6090267	Minus	13714-15440
402368	1506866	Plus	4426-4646
402469	9797107	Minus	71266-72351
402532	9630951	Minus	162240-160586
402560	6644273	Plus	33839-33715
402575	6684830	Minus	109742-109883
402602	7236666	Plus	6786-6972,7478-7575
402756	9713669	Plus	67838-87924
402786	9715046	Plus	47624-47765
402807	6466148	Minus	101642-101660,103476-103656
402810	6010110	Plus	12715-12866,13627-13643
402864	5561599	Minus	46824-46784
403046	3640163	Minus	56707-66956,66369-66611
403055	6748904	Minus	109532-110226
403217	7630969	Plus	54089-54163,55427-55623
403218	7630969	Plus	58039-58149
403291	7230970	Plus	56177-56436
403339	8448936	Minus	120428-120703
403554	6736083	Minus	28634-28756
403704	4892546	Minus	6850-6966
403708	5725381	Minus	134394-134612
403726	7534331	Plus	63737-68843
403739	7530892	Plus	44563-44766,48203-48483,52255-52405
403740	7530882	Plus	88504-87227
403742	7552036	Minus	67610-68002
403748	7552036	Plus	59612-59897
403885	7710403	Minus	53259-53504
403893	7710691	Minus	5435-7846
403947	7711523	Plus	38557-38617
404036	6698763	Plus	61809-62011
404034	3548795	Plus	68713-69175
404058	3548795	Plus	99387-101808
404108	6207474	Minus	63603-64642
404211	5006246	Plus	186728-185985,194575-194688
404277	1534458	Minus	91855-91946
404364	8897028	Minus	38055-38156,42175-42381,43435-43553
404407	7329316	Minus	48154-48469
404489	8113772	Plus	98183-98489
404527	8152097	Plus	127737-127796,128080-128210,129688-130054,132546-132889





WO 02/098358

PCT/US02/17594

Table 3A shows the Seq ID No, Pkey, ExAcon, UnigeneID, and Unigene Title for all of the sequences in Table 4.

Pkey: Unique Ecol probe/ identifier number

ExAcon: Exemplar Accession number, Genbank accession number

UnigeneID: Unigene number

Unigene Title: Unigene gene title

Seq ID No: Seq ID number correlation for those sequences in Table 4

	Pkey	ExAcon	UnigeneID	Unigene Title	Seq ID No
	415539	AF33981	Hs.72772	BMP-RIE	Seq ID No 1 & 2
	448988	Y09783	Hs.22785	gamma-aminobutyric acid (GABA) A receptor	Seq ID No 3-10
	403740			NM_001076*Homo sapiens UDP glycosyltran	Seq ID No 11 & 12
	408633	AF563372	Hs.46677	PROR200 protein	Seq ID No 13 & 14
	408650	AA525775		ESTs, Moderately similar to PC4259 (nm)	Seq ID No 15 & 16
	409051	AA08912		gb:aa6403.r1 Stratagene hNT neuron (B37	Seq ID No 17
	409123	AA063403		gb:zn0412.s1 Stratagene corneal stroma	Seq ID No 18
	415787	H01463	Hs.93534	ESTs	Seq ID No 19-21
	415999	AA172179	Hs.294025	ESTs	Seq ID No 22
	416225	AA577730	Hs.185684	ESTs, Weakly similar to PC4259 (nt)in	Seq ID No 23
	420757	XT8592	Hs.99915	androgen receptor (dihydrotestosterone r	Seq ID No 24 & 25
	429163	AA884766		gb:aa20a10.s1 Scores_NFL_T_GBC_S1 Homo s	Seq ID No 26
	429441	AJ224172	Hs.204036	Ippophilin B (uteroglycolin family member)	Seq ID No 27 & 28
	431099	Y13387	Hs.249235	phosphoenolpyruvate-3-kinase, class 2, alpha	Seq ID No 29 & 30
	432432	AA541323	Hs.115931	ESTs	Seq ID No 31
	432435	BE218996	Hs.282070	ESTs	Seq ID No 32 & 33
	432527	AF675028	Hs.102754	ESTs	Seq ID No 34
	435876	AF612586	Hs.160271	G protein-coupled receptor 48	Seq ID No 35 & 36
	436233	W52448	Hs.56147	ESTs	Seq ID No 37-40
	436569	AF602166	Hs.222399	CEGF1 protein	Seq ID No 41 & 42
	440819	AF094644	Hs.202108	ESTs	Seq ID No 43
	442832	AF206590	Hs.253569	ESTs	Seq ID No 44
	447342	AF190268	Hs.191392	Homo sapiens, Similar to RIKEN cDNA 2010	Seq ID No 45 & 46
	447499	AF262580	Hs.147674	protocadherin beta 16	Seq ID No 47 & 48
	451411	AA017492	Hs.135655	EST	Seq ID No 49
	451720	AF070885	Hs.290853	ESTs	Seq ID No 50 & 51

WO 02/098358

PCT/US02/17594

Table 3B shows the accession numbers for these Play's lacking UnigeneID's for table 3A. For each probe set is listed gene cluster number from which oligonucleotides were designed. Gene clusters were compiled using sequences derived from Genbank ESTs and mRNAs. These sequences were clustered based on sequence similarity using Clustering and Alignment Tools (DoubleTwist, Oakland California). Genbank accession numbers for sequences comprising each cluster are listed in the "Accession" column.

5	Play	CAT Number	Accession
	408660	107794_1	AA025775 AA056542 AJ338978 AW975281 AA664986
	409051	109690_1	AA080912 AA075318 AA063403 AA076594 AA078992 AA084826 AA081881 AA113913 AA113892 AA033821 AA134801 AA082953 AA070343
			AA062835 AA075419 AA063293 AA071252 AA078900 AA062836 AW974305
10	409123	110143_1	AA063403 AA070823 AA070950
	429153	300543_1	AA084765 AW974271 AA062975 AA447312

WO 02/098358

PCT/US02/17594

Table 3C shows genomic positioning for those Pkay's lacking Unigene IDs and accession numbers in table 3A. For each predicted exon is listed genomic sequence source used for prediction. Nucleotide locations of each predicted exon are also listed.

	Pkay	Ref	Strand	NT position
5	403740	7630862	Plus	66504-67227

WO 02/098358

PCT/US02/17594

Table 4:

Seq ID NO: 1 DNA sequence

Nucleic Acid Accession #: NM\_001203

Coding sequence: 274..1782

	1	11	21	31	41	51	
	1	11	21	31	41	51	
	GGCCGCGGCG	GCATCGCGCG	GGGCTCGCGG	GGAGCGCGCG	AGTGGCGAGA	CTCGCGCGCT	60
10	GAGGACGCGG	GAGCGCGGAG	GGCAGCGCGG	GGGTGGAGTT	CAGCTACTCT	TTTCTAGAT	120
	GTGAAGAGAA	AGGAAGATCA	TTTCACTGCT	TGTTTGAATA	GTTTCAGAGT	CTTCTGAGCT	180
	CTATACATT	TGCTTCGAG	CTATGCAAG	AGAGAGAAAC	AAAGATTAA	CTTCAAGGCG	240
	TGATTAAGT	GGAAAGAA	CTTCTCGAT	TACATCTT	TGGAAATTA	AGGAAATTA	300
	AATCTGGCA	CTGAAGAAAC	GGATGTGAG	AGTACAGCGC	CCACGCGCGC	TCCAAAGGTC	360
15	TTGCTGTGA	AATGCAACCA	CCATTTGCCA	GAAGACTCAG	TGAAATATAT	TTGCAACACA	420
	GACGCAATAT	GTTCACAGAT	GATAGAGAG	GATGACTCTG	GGTTCGCTGT	GGTCACTTCT	480
	GGTTCCTGAG	CTGATGAGG	CTGAGTTTT	CAATGTGCGG	AGACTCCGCT	TCTCTATGCA	540
	AGAAATCAA	TGATATCTG	CACAGAAAGG	ACCAATATTA	ATAAGACCT	ACACCTTACA	600
	CTGCTCCAT	TGAAGAAACG	AGATTTTGTT	GATGAGCTTA	TACACACAG	GGCTTTACTT	660
20	ATATCTGTGA	CTCTCTGTAG	TTTGCTCTTG	GTCTTTATCA	TATATTTTGT	TTACTTCGCG	720
	TATAAAGAC	AAGAAACCG	AGCTGATAC	AGCTATGGGT	TGAAGAGGA	TGAAGACTAC	780
	ATTCTCTGT	GAGATATCT	CGAGACTTA	ATGAGAGCT	CTCAGAGCT	AGGATATGA	840
	TCAGCTCCT	CTCTCTGTT	CCAAAGGACT	ATAGCTAAGC	AGATTACAGT	GGTGAACAG	900
	ATTGGAAGA	GTGCTATGG	GGAGTTTGG	ATGGGAAGT	GGCTTGGCGA	AAAGTATGCT	960
25	CTGAAGTTT	TCTCTACAC	AGAGCAACC	AGCTCTTCA	GAAGAGGGA	ATAATATCAG	1020
	ACAGTTTGA	TGGAGCGA	AACTCTTTT	GGTTTCTATT	CTCGAGATTA	CAAGAGGCA	1080
	GGGTCTTGA	CCGATTTGA	CTTAATACA	GACTATCAT	AAATGTGTC	CTTTATGAT	1140
	TATCTGAAT	CCCAACCT	AGACCTTAA	TCATTTCTGA	AGTTAGCTTA	CTCTCTCTCT	1200
30	ATCTGTTAT	GTCTATTGA	CCAGAAATC	TTTATGACT	AGGCAACCC	AGCACTGCC	1260
	CTATGAGCA	TGAATAATTA	AACTCTCTG	CTGAGAGAA	ATGAGAGCT	CTATATCTCT	1320
	GACCTGGCC	TGGCTTTAA	ACTTATAGT	GATACAAAT	AAATGTGAT	ACCACTTAC	1380
	ACTGACTTG	GCACCAAGC	CTATATGCT	CCAGAGGTT	TGAGAGGAG	CTTGAACAG	1440
	AATCACTTC	AGTCTTACAT	CATGCGTAC	ATGTAATATT	TGCTCCGAT	CGTTTGGGAG	1500
35	GTCTTGA	GATGATGAT	ATGAGAGTA	CTGAGAGTA	CTGAGAGTA	CTGAGAGTA	1560
	CTATGTC	CTGACCTCT	TTATGAGGAC	ATGAGAGGAG	TTGTTGSCAT	CAAGAGTAA	1620
	CGCCCTCAT	TCCCAACCG	GTGAGGCA	GATGAGGTC	TAGGAGGAT	GGGAAAGCT	1680
	ATGACGAAT	GTGAGGCGA	CAATCTTGA	TCAGAGGCTA	CAGCTCTGCG	GGTTAAGAAA	1740
	CACTCTTCA	AAATGAGTA	CTGAGAGG	CTGAGAGG	CAAGAGTAA	GAAGATTAAG	1800
40	CACTCTTCA	GAAGGCGAC	AGCTTACTT	CTGTTTGTG	GAGAGGAA	AGACATCAA	1860
	TAGAGTCCA	CAATCAAGC	CTGAGAGCT	GTCTCTGCTT	CGATGTGGTT	CAGAGCTCAC	1920
	CTTCAAGGA	GCGACTGCG	CAAGAGAGA	GAAGTCTCA	GAAGAGAGA	TTGATCTCTG	1980
	TCCTTTTGA	GGGAGAGAA	CGTTGGGCTA	ACTTTGTTCA	GATATGATGC	AT	

Seq ID NO: 2 Protein sequence

Protein Accession #: NP\_001194

	1	11	21	31	41	51	
	1	11	21	31	41	51	
	MLLESAGKLN	VGTKKEDGSS	TAPTPRPKVL	RCKCHHHCP	DEVNHCSTD	GYCFMIBED	60
50	DSLPVPTSS	CLGLBSDFQ	GGTFIPHGR	RSIBCTERN	ENGLDPLFL	PLINDRFPD	120
	GPIMHRLAL	RYVTKSLLY	LTLFCTPFI	KGRTFPFIS	TGLRQDTYI	PPISRLRLI	180
	RQSSGSSGSS	GLPLVQRTI	AKQIQMVQI	CKRYGVPMH	GMWRGKVAV	KVFTTSEAS	240
	WPRETIYQT	VIMRHENGLG	FIADIDKGTG	SWGLYLITD	YHEMGLYSY	LKSTLLDAS	300
55	MLRLAYSSS	GLGLRTEIFP	STQKRIALH	RLDSRLMLV	IKNGFTCLAD	LGIAVFTSD	360
	THVDIPFMT	RVTGSRVMP	FTLBSRLAH	TPSFLTDLH	YFGLILHIV	ASRCGSDIV	420
	REYQLPHELT	VPSPDSYDM	REIVCTKKLR	PSFPRWSSD	ECRLQMKLM	TECMANFAS	480
	RLTALRVKKT	LAMSSQSDI	KL				

Seq ID NO: 3 DNA sequence

Nucleic Acid Accession #: NM\_004961.2

Coding sequence: 55..1575

	1	11	21	31	41	51	
	1	11	21	31	41	51	
	GCGAGAGCGT	GAGCGCGGAC	CTCGCGCGAG	GTGCTCGCGC	CGCTCTCGCG	GGAATATGTT	60
65	TGCAAGATTC	TTGCACTGCT	CTGAGGATC	TTATGTATGC	TGCACTCGAG	GGTCAAGGGA	120
	CTTGAGATCT	ATATCAAGAA	TGAGGCTTAT	TGCTGATAT	CTGCTGATG	CCGCGAGCG	180
	CAGCTCTGCT	AAATACAGCT	CTCTCTGAG	GAACAAAGT	CAACTAGAGC	TGAGCTCGG	240
	AGCAGCATTT	GCAAACTGCC	AGAGGCTCT	CGCATCTGTA	ACATCTATCT	GATTAATAT	300
70	GACCAAGAAC	TGGCGCTGCG	CATTGAGAG	AAGCCCACTG	TGGTCATGTT	TGAGATCGCC	360
	GTGACAGGCT	TGCTCTGCTT	CTCTATCTTA	GACATGAGAT	AGCATCTGTA	CACTGCTCTC	420
	TGCGAGACT	GATACAGGCA	AGCTCTCTCT	TACAGAGCA	CTTCTATGCT	TCTTCTCTG	480
	AATCCCAATG	TGCTTGAGCA	GCTATGGATC	CGGACACCT	TTTTTAGGAA	TCTTAAGAGG	540
	ACCGAGAGC	ATGAGATGAT	CATGCCCAAC	CAGATGCTGC	CGATCTACGA	GGATGGGAG	600
	GGTCTGATCA	CAATCAAGAT	GAGACTTAT	CTGCTCTGAT	CTTCTCTGAT	CTTCTCTGAT	660
75	CCGATGATTT	CTGCTCTGCT	CGCTCTATCT	CTCTCTGAGT	TTTCTCTATC	GGAGAGGAG	720
	ATGATCTACA	AGTGGGAAAA	TCTCAAGCTT	GAATCAATGT	AGAGGAAGCT	CTGAGAGCTC	780
	TTTCAAGTTT	ATTTCAGAGC	AGTGAAGAAC	AAAGCTGAAA	TAACTACAC	CCGAGTTGGT	840
	GACTCTAGG	TGCAAGAGG	TTCTCTCTCT	ATGAGCTGAG	TGCTTGCTTA	TGTGCTGAG	900
80	CAAACTATCT	TGCTCTCTCT	CTGAGAGCT	ATGAGCTGAG	GGTCTGCTTA	TGTGCTGAG	960
	ACAGAGCTCT	CTCGAGCGCG	GAGCTCTCTA	GGATCACTCT	CTCTCTGAT	CATGAGCAGC	1020
	TTGCGGCACT	TTTCTCTCTA	GAATTTCTCT	CGTCTCTCTT	ATATACAGC	CTTGATATTC	1080
	TATATGCGCA	CTCTCTCTCT	CTCTCTCTCT	CTCTCTCTCT	TGAGATTTGC	TGCTCTCTAC	1140

WO 02/098358

PCT/US02/17594

5  
 10  
 15  
 25  
 30  
 35

TTCTCATCT ACAACGAGC AAAGCCCAT GCTTCTCTTA AACTCCGCCA TCTCTGTATC 1200  
 AATAGCCGTG CCAATGCCCG TACCCTGTGA GCTTCCCGAG CCGTGTGCCG CCAAGATCAG 1260  
 GAAGCTTTTG TGTGCGAAT TTTCATCATT CAGGAGAGTG ATGAGGAGGA GCGCCCGTCT 1320  
 TGTGCGAGC AGCAGCGAGT TACGACGAGT AGCCGCGAGG GTCCCGCGAG CTTCTCTCTC 1380  
 AAGCTGCGCT GCTGCGAGTG GTGCGAGGTT TTAAAGAAT ACTCTGCACT GGTCCCGCAT 1440  
 TGTGAGGCGA GTACTGTGCA GCGAGGCGCG CTTCTCATCC ATGTCTACCG CTTGCGATAA 1500  
 TACGAGGAG TTGTTTTCGC AGTACGCTTC TCTCTCTTA ATGTGCGCIA CTGCGTTGTC 1560  
 TGTCTTAACT TGTAGTGTGC TGTGCTGAGC CTGTGAGGCA CTTCTCTGCA GTTCCGAGA 1620  
 GTTCCAGAGC CTTCTCTCAG GAGTGTGCGG GAAGCGACGA CAGCAGCATG GACCGCTAG 1680  
 AGTTTTCCTC GCGCCATTC CCAACAGGAA GCTTGCAGAG GGTTTGTCTT TGTGCGCCCT 1740  
 CTCCCGTACC TGGCGCATCT ACTGAGTCTT CTGACGAGC CATTTCMAAT TATTAAATAA 1800  
 TGGCCGACCT CCGTCTCTTT GAGGAGCAT CCGTGTGACT AGTGTTCGAA AACGACGAGC 1860  
 ACTTACTGAT GAGTCTGCTA AACCAAGGCT TAAGTACAGC GATTACTACT ACTTCCAG 1920  
 AATGCTGACC ACCAGAGCAT TACTGCTATT TTCCAGAGC CCACTATTGC CTTGTAGTGT 1980  
 CTTTGGCCCC AGTTCTGGCC TCAGCCTGTA AGTCACCGA CTAATGCTCT GCTATACACT 2040  
 GCGCCGCTAT TAAAGTCGTG GCGACGATTA TAAACAGAGG AAGAGATGCC TTCTCTTTTG 2100  
 TCAGATGAT ATGTGTCGAG TCTCTCTGCT CCGTACGCC TTCTCTGCA GTATGATAGA 2160  
 CAGTGCATTT ATGCTTTTAA GAGAGCGGCG GCGCACAG AGAGCGCTATT TGGGACAGA 2220  
 TTCTCTCTCT TCTGCTGCTG TGACATCTCC CTTCTCTTTC TGTGCTCATC TTTGCTCTGC 2280  
 ACTACCATTT CAATGCGCTT CATCAATGCG GTATCTGTT TTGTGTGTGA TTACATGATC 2340  
 TATGCTGCT TTATGATGCG ACCCTTTCTC TCTCTCTGTA CCGTGTGAG TCTGTGCTTA 2400  
 ACTTCTCCAG TCACTTCCGC TACGCTGAGC CAGCGACTTA GCGCTTGGTG ACTTCTGGG 2460  
 CGCAAGAAC TAAAGAACT CCGCTTTGCA ACAGCGATTA CTGCGCATTG ATTGTGCCC 2520  
 ACCGAGGCGA CACTGTGCGA GTCTATTCAC TTCTGTAGC CTTGCGGCCA TAAACGATC 2580  
 CAGCTGTATA CCGCGCGGAC TCTAGACATC ACAGCATC ATCTAAATTC CTTTAAATTT 2640  
 GTATGGCACT GGAACCTTGG CAAAGCACTT TTGCAAGTT GTGTCTGATT GAGCTTCAT 2700  
 GATAGCTGT TGACTGTCTT AGGCGAGCAT TCTTATCCCG ATTTCGAGA TGAAGAACCT 2760  
 GAGTCACACA TTCTGTGGG AGCTGTGAGT TCATCTGAGG CATCTCAGA GCGCCATCTG 2820  
 ACTCTGAGA CGATGATAGA CAGTCTGCTA CCGATGATCT CTTCTCTGAC TCTGTCTG 2880  
 CTTCTCTGCG ACACCGTGG CAGGCGCGAG AATGCGAGCC TCTCTTAGC TCAATTTCTG 2940  
 GCGCTGAGGT GCTGAGACTG CCGCCAGAT CAAATCTCTC CTGCGCTGAG TACCCGAGTG 3000  
 GATGAGACTT GAGCATGCGC CAATGCTCTT ATATGCTTAG TGAATCTGT GTCTGTATTT 3060  
 TCTTGTGCGG TGTGATGTA GAGTCTCTCC TCTCTTCTTT GTTCAACATC TGTGAAATGG 3120  
 GGAATATGCT AATCAATCAT ACCGCAAG ACAGAGAAA AAAAAAAA

Seq ID NO: 4 Protein sequence  
Protein Accession #: NP\_064952.1

40  
 45  
 50

1 11 21 31 41 51  
 LSKVLVLL GILLIQGV EDPETESME ASSEGVVTF QPPLBNLL SEETSTETE 60  
 TSSVRKLPRL ASRIANTILS NYDKLRFPI GEPFTVTV LARSLGPSL FLDMETCTDI 120  
 IFSGWYDER LQNDTFEEL VLAMVVVQL WIFDTFFNS KRTHBEHITM FQHVRIYED 180  
 GKVLVITRM IDAGCSLML RFPMDSBSC LSFSSFSVPE NMYIKHNF KLEINEKISN 240  
 KLFDFPTFV SNTSLITP VQDFVWITF RVNRSFVGV AFQWVPSVS TMLSWPSFN 300  
 HTBSRAFT SLATSVLVN TLGTFGRM FFFVSTTAL DPTIALCFV CFALLERAY 360  
 LKFLVYNQK AHASRKLHP RINSRAHAK RFRKACARQ HQBAPVQIV TTBGSGDER 420  
 PSCBAQQPS PGPBPGRSL CSKLACCNC KRFKKVFCMV PDCBGTWQQ QRLCIEVRL 480  
 DNYSRVVFV TFFFFSVLVN LVCLML

Seq ID NO: 5 DNA sequence  
Nucleic Acid Accession #: NM\_021984.1  
Coding sequence: 572..1753

55  
 60  
 65  
 70  
 75  
 80

1 11 21 31 41 51  
 GCGAGAGCT GAGCGCGAG CTTCCGCGAG GTGCTGCGC GATCTCCG GGAATATTTG 60  
 TCGAAGTCT TTGCTGCTC CTTGCGCTCT TTATGTAGC TCGATCGA AACATGTATA 120  
 CAGAGATGTG CTCATATCAT ATGTGTACAG CTGATGAT GTGAAAAAT GACCAAGCG 180  
 GTGTAAAGAA AGCAAAATCA AGGACCGGAA TGTGACAGG ACGTCAGAG CCGCTTTTGT 240  
 CAGTGGCTCC CAGCAAGAGC AGCATATCC GATCTCTTA CAGCATCGG TCGAGAGGCC 300  
 TCGACGTA TCGGCTGTG AGCTCTCTCT CAGTCTATC CTTGCTGCTT CCGAGGCCA 360  
 GCGCTCGGAA AATGAGCTCT TCTCTGAGGA CAAAGATCA ACTGAGACT AGATCGGGAG 420  
 CAGAGTGGC AAATCGCCAG AGGCTCTCG CATTCTGAC ACTATCTGA GTAATTAAGA 480  
 CCAAGAACT CCGCTGTGCA TTGAGAGCA GCGACTGTG GTGATCTTTG AGATCTCGT 540  
 CAGAGCTCT CCGCTCTCT CCGTCCGGA CTGAGACAG ACATATGCA TGTCTCTTC 600  
 CAGAGCTGG TACAGGAGAC GCGTCTGTTA CAGAGACAC TTGAGTCTC TTCTCTGAA 660  
 TGGCAATGT GTGAGCCAGC TATGATGCC GGAACCTTT TTGAAGATT CTAAGAGGAC 720  
 CCAAGACAT GACATACAGA TCGCCACCA GATGTCCCG ATTCACAGG ATGCGAAGT 780  
 GTTACACAGA ATTAGAGTA CATTATGCT CAGTACCTCA CTCACATG TGAATTTTCC 840  
 ACTGATGAT CACTATCTCT CTTCTCTCT CTTGCTGAT TCTGATGCT CCGAGGCCA 900  
 GTTCAACAG TCGCAAAAT TCAAGCTTGA AATCAATGAG AAGACTCTC GAGAGCTCT 960  
 CAGTGTGAT TTACAGGAG TGAAGAACAA AACTGAATA ATCAACACC CAGTTGTGTA 1020  
 TCTCATGTC ATAGATGCT TTCTCAATG GAGCAGGCG TTTGGCTATG TTGCTCTCTA 1080  
 AATCATGCT CCGTCTCTCT TGAACAGAT CAGTCTGCTT GTTCTCTCTT GGTATCAGC 1140  
 AGAGTCTCT CAGGCGCGA CTTCTCGAG GATCACTCT GTTCTAACCA TGACAGCTT 1200  
 GCGCACTTT TCTGTGAAGA ATTTCCCGCG TGTCTCTAT ATCAAGGCT TGAATTTCTA 1260  
 TATGCGATC TGTCTGTCTT TGTCTGTGTC GAGTGTGTC TGTCTGTGTC TGTCTGTCT 1320  
 CTTATCTAC AACGAGCA AGGCTCATG CTTCTCTTAA TGTCTCTTAA TCTGTATGA 1380  
 TATGCTGCC CAGTCCGTA CCGTGTGAG TCTCGAGCC TGTGCGGCC AACATCAGA 1440  
 AGCTTTTGTG TCGCAGATTG TCACACTGA GGAAGTGTAT GGAGAGGAGC GCGCTCTTGT 1500  
 CTCAGCCAG CAGCTCCGTA CCGGAGTAT CCGTAGGCT TCGCGAGCC TCTGCTCTGA 1560  
 GCTGCGCTCC TGTAGTGAT CAGAGCTTT TACAGATAC TCTGCGATG TCCCGCATG 1620



WO 02/098358

PCT/US02/17594

5  
10  
15  
20

GCCGACTATT GCGTTTGGAG TCGTTTGGCG CAGATCTCTGG CCTCAGCCTC AAGAGTGACCC 2100  
GACTAGTTCG TTGCGTATAC CTGGCACTTC KTTTAAKTCG TGGCGAGGAG TATAACGAGA 2160  
GAGAGAGATC CCGTCGCTTT GCGCAGATTA TTAATGTTCT AGTTCTCTCT CCCGCTCATC 2220  
CGTTTCTCTG CAGATACATA GACACGTGGA TTATCGCTTT AGAAGAGGAG GGGGGCAGCA 2280  
AGAGAGATCA TGGTGGAGAG GATCTCTGCT TCTCTCTCT TGGGAGCTCT CCGTCTCTCT 2340  
GTGCGATCCA TCTTTGCTCT GCGATACCAA TCGAGTGGCC TTCAACCAAT GGGTATCTAT 2400  
TTTGTCTCTG GATTATAGTA ACTACTCCCT GCTTTATATG CCAAGCTCTT CTTCTCTCTT 2460  
GACCGCTCTG ACTCTCTCTG TAACTTTCCG AGTACATCTC CTTAGCCCTG ACCAGAGCACT 2520  
AGCTCTCTCT TGGTGGAGAG CTAAGAGGAG TCGCTCTCTG TCGCTCTCTG CAGAGCTCT 2580  
10 GACTCGCTCT GATTTGCTGG CACCGGCGCC AGTCTATATC CTTCTCTCTG CTTCTCTCTG 2640  
CCCTGAGCCC ATAAACCGGT CCACTCTTAT ACCCGGGGCA CTCATACCAT CACATATCAT 2700  
CAATCAAAAT CCGTTAAAT TGTATGGGAC TGGACCTTTG GGAAGGACAT TTATGACAGT 2760  
CTCTCTCTCT TGGATCTCTA CTGATCTCTG CTGATCTCTG TGGATCTCTG TGGATCTCTG 2820  
CACTCTCTCT ATGAAACCGC TGGATCTCTG ATTCTCTCTG GACTCTCTCTG CTTCTCTCTG 2880  
15 GCTATCTCTG AGCCGACTCT CACTCTCTCT GACCATCTCT AGGCTCTCTG AGCTCTCTCT 2940  
ACGATCTCTG TCTCTCTCTA CTTCTCTCTG CACCATCTCT GCGAGGCGCA GATGCGGACAT 3000  
CTCTCTCTCT CTCAATCTCT GCGCGCGGAG TCGTCTCTCT GCGCGCGGAG TCAATCTCTCT 3060  
CTGCGGCA CTGATCTCTG GGTATGAAAT TCGCTCTCTG CCGATCTCTG TTATCTCTCT 3120  
20 GTGAAATCTG TGTCTCTCT TGTGTTGGG GTGATGAGG TGGGCTCTCT ATCTACTCTT 3180  
TGTCACTCTG ACTGTAATG GGAATATG TAATAAATA TATCGCAAA GC

Seq ID NO: 8 Protein sequence  
Protein Accession #: NP\_068822.1

25  
30

1 11 21 31 41 51  
MEYTDIIFS QTNWKEKTHE HEITPMQWV RIYXGKVLV TIRMTIDAGC ELHMGRFMD 60  
SSSCPLPSS FSYPEREMII KWNPKLEIN BKISMKLPF DFTVSHKTS LIITFVDFMD 120  
30 WTTFPNSVR FQVYAPQYI VPSVYVXLS WSPVSKTBS ABPRTSLGIT SVLHTLTAT 180  
FRRKSPST IYTAAGATA ICFATPCAL LEFWLFIWLI YNGTARABR YNGTARABR 240  
ABATRASER ACARQHQEAT VCQIVTEBS DGEERPECSA QQPFQSPSE PPELSCSKLA 300  
CCEWRKFRK YFCWQHQEAT STWQGRGLCI HVYRLDNYSR VVFFVFFVF NVLYWLWLCNL 360  
L

Seq ID NO: 9 DNA sequence  
Nucleic Acid Accession #: NM\_021990.1  
Coding sequence: 1309..2490

35  
40  
45  
50  
55  
60  
65  
70  
75  
80

1 11 21 31 41 51  
GCGAGAGGCT GAGCGGCGAC CTCGCGCGAG GTGGTCCGCG CGGTCTCGCG GGAATATGTT 60  
TCGAGAGTTC TTGCGTCTCT CTTAGGCGAC TTATATGATC TCGATCTGAG AAGATATATA 120  
GAGAGAGTGS CTGAAATCTG AGGTGTACAG CTGATGATGT GTCAAAAGAT GCGACAGCT 180  
GTGTAAAGAA AGCAAAATCA AGGACCGGAA TGTGACAGAG ACGTCAGAGG CCGCTTTGTT 240  
45 CACTGCGCTCG CAGCAAAAGC AGCACTATCC GCACTCTCTA CACCATCTGT GATTTCTCAT 300  
CTCTGAGAGA TGGCTCTTAA CATTCTCTCT TAATCTCTCT ATTCTCTCTA ATCTCTCTCT 360  
TTTCTCTGCG TGTGTTGAT CTTCTCTGAG CTGAGGTGCG ACTCTCTCTG CCGACCGCAT 420  
TCTGAGTGTG CTGCTCTGTT GTGAGAGACT CCGTCTCTCT CTGTTAGAC ACCCAAGAG 480  
50 GCTGCTCTCT AGCTCTCTCT CTTCTCTCTC CTTCTCTCTG CCGGAGTCTA ACAGATATTA 540  
CGACAGCAAC AAGACCGGAA AATATCTCTA CAACTTTCTG GTCTCTCTCT GAGAGAGGCG 600  
CTCTGCTCTG TCTCTCTGAG CCGGCTCTCT TCTCTCTCTG TCACTCTCTG TGTGCTCTG 660  
TCAGGCTGAG TAGAGCGCAA GCGGACCAAC ACTAGGCAAA CCGCGCCAGC CCTCAGACAT 720  
AAATGCGCTCT TTCACTCTCA GTGTAACTAT CTTTAAATCT CTAGGCTCTG GTTTTGTGTA 780  
45 TTTTCTCTTA ACTAAAGAGT TGAATATATA AGGACGACAG CATAGAAAGT CCGCAAGAG 840  
CGACAGAGCT TCAAGAGGCT APTCTCTCTG TCTCTCTCTG TCTCTCTCTA TCTCTCTCTA 900  
CCAGTCAAAA TTAACTAGG TTTTGTGTTG AAACTCTGTT TGTGTTGCT TCTCTCTGAA 960  
GAGGACCTAG CTGGGCGCCC TACAGAGTGC AGGGCGAGAG TTCACTTTCT GTTTGAAATGT 1020  
50 TCTAGAGTGC AGGACACTCA GACTGATCTA AAGATAGAG CCGCTCTGCG TATGTTGTTCT 1080  
TATCGCCCC AGCCCGGCG TCTGGAAGAT GACTCTCTCT CTGAGAGAG AAGTCTCACT 1140  
60 GAGACTGAGA CTGGGAGGAG AGTTGCGAAA CTCGCGAGAG CCGTCTGCAAT CTGGAACACT 1200  
ATCTGAGTGA ATGATGAGCA CAATCTGCG CCGTCTGATG GAGAGAGGCC CACTTGTGTT 1260  
ACTGTTGAGA TCTCTCTGTA CAGCTCTGSP CCGTCTCTCT TCTGAGACT GAGATAGACC 1320  
ATGATCTCA TCTCTCTGTA CAGCTCTGAG GACGACAGCC TCTGTTAGCA GAGACTCTT 1380  
GAGTCTCTG TCTGTAATG CAACTGTGTT AGCCAGCTAT GATCTCTGGA CACTTTTTT 1440  
65 ACCAATCTTA AGAGAGCCCA CGAGCATGAG ATCAACATCG CCAACAGAT GGTCCGCGATC 1500  
TACAGAGTGG GCAAGTCTGT GTTACAAAT AGGATGACCA TTGAGTCCGAG ATGCTCACTC 1560  
CAAGTCTCTCA GATCTCTGAT GATCTCTGAC CAGTCTCTCT TACTCTCTCT TACTCTCTCT 1620  
TACTCTCTCA ATGACAAAT TCTGAGTGGT GAAATTTTCA AGCTGAAAT CTGAGAGTGA 1680  
AACTCTCTCA AGCTCTCTCA GTTCTGATTT ACAGAGTGTG GCGACAAAG TGAATATATC 1740  
70 ACAAACCGAG TTGAGTACTT CATGCTGTGT ACAATTTTCT TGTATGTGAG CAGGCTGTTT 1800  
GGTATCTGTT GCTTCTCAAA TATGTCCTCT TCTCTCTGTA CAGAGTCTCT CTTCTGATTT 1860  
TCTTCTGTA TCTGAGTGA GTCTCTCTCA GCGGAGACT CTCTCTGAGT CAGCTCTCTT 1920  
CTGACCTGTA CCACTGTTGG CAGCTTTTCT GGTAAAGAA TCCCGGTTG TCTCTCTCTA 1980  
ACAGCTCTCTG ATTTCTATA GCGCATCTCG TTCGTCTCTT CTTCTCTGCG TCTGTTGAG 2040  
75 TTTGCTCTCA GTATCTCTCT GATCTAGAC CAGGACAAAG CCACTCTCTC TCTTAAATCT 2100  
CGCATCTCTG GTATCTCTCT CAGCTCTCTG GACTCTCTCT CTGAGCTCTG CAGCTCTCT 2160  
GCCCGCAAC ATCAGAGAGC TTTTGTGTTG CAGATGTGTA CCGTCTGAGG AATGTATGTA 2220  
GAGAGGCGCC CGTCTCTCTC AGCCGAGAG CCGCTCTGAG CAGTATGCC TGAAGTCTCC 2280  
80 GCGAGCTCTG CTCTCAAGCT GCGCTCTCTT GAGTGTGTA CCACTCTCTC TCTTAAATCT 2340  
TACGATCTCT GTCTCTCTCT CAGCTCTCTC TGGCAGAGG GCGAGTCTCT TCAAGTCTCT 2400  
TACGCTCTG ATACTACTCT GAGGCTGTTT TTTCTCTGTA CTTTCTCTT CTCTAATGTT 2460  
CTCACTCTG TGTCTCTCT TAACTGTGAG GTACCGAGTG GTACCTCTCT GAGGCACTCT 2520  
TCAAGTCTCC CAGAGGCTCT AAGCGCTCTT CCGAGGGAGT TGGGGGAAAG CAGGAGGAGC 2580

WO 02/098358

PCT/US02/17594

ACGAGGAGCG ACTAGAGTTT TTGCTGCCCC ATTCGCCAAA CAGAGAGCTG CAGAGCGTTT 2640  
GCTCTTCTCG CCCCCTCCCC CTACCTGGCC CATTCACGTA GTTCTCTCAG CAGACCATTT 2700  
CAAAATTATTA ATAAATGGCG CACCTCCCTC TCTCTCAAGG AGCATCCGTC ATGCTCATGT 2760  
TTCPAAACCA CAGCAGACTA GTGATGACT CCTTAAACAC ATGCTCAAGT ACAGGCGAAT 2820  
TAACTATCTG CAGAGACTGT GAGATATCTG CATTTTCTCA GAGAGCGACT 2880  
ATTCCCTTTG CAGTCTTTCT GGCCAGATTC TGGCCTCAGC CTAAAGTCAG ACCGACTAGT 2940  
TCCCTGGCTA TACCTGGCAC CTCATTAGA TCGTGGGCGC CAGTATAACA GAGGAGAGAG 3000  
CACTCCTCTC TTGAGTCAGA TTATATAGTT CTGAGTCTCT TCTCCGCTGT ACCCCTTTCT 3060  
CTGAGATGAG AATAGACTGT GATATATCCC TTATGAGAGA GGGGGGGGCA CAGAGAGACT 3120  
CTCATTTGCA CAGCATCTCT CTCTCTCGGC TGTGAGACAA TCTCCTCTCT CTCTCGGCTC 3180  
CCATCTTCTG TCTGCACTAC CAATTCAATG CCTCTCTATC AATGGGTATC TATTTTTGTG 3240  
TCTGATATTA TGAATACTCT CTGCTCTTAT ATGCGACCCCT CTCTCTCTCT TCTGCGCCT 3300  
GTGACTCTTT CTGATAGTTT TCCCTGACTC CTCTCTCAGC CAGATCAGC ACTATGCTCT 3360  
GTGCACTCTC TGGGCTCAGG AATCAATAGA AACTGCGCTT TGGACAGGCG ATTAATCTCC 3420  
ATTCACTGCT GCGCAACGAG GGCACACTGT CGGAGTCTTA TCACTTGCTT GACCCCTGGA 3480  
CCGATCAACC AGTCACTGTT TATACCGGAG GCACTCTAAC CATCAATCAT AATCAATCAA 3540  
ATTCCCTTAA ATTTTATATG CAGTCGAGAT TTGCGAAGC ACTTTTGA CA ATGCTGCTCT 3600  
GATGAGACT TCGATATAGC CTTGTGCTGT GTTGAAGCA GATTTCTTA CCCCTTTTG 3660  
CAGATGAAAC CCTGAGTCA CAGATTCTGT TGGAGCTCTG GATCTCACTG GAAGCTATCC 3720  
AAGAGCCGAC TGTGACCTTC TAGACCAAT GATAGGGCTA GAGGCTCAG TCTCACATGA 3780  
TCTCTCTCTG TCACTCTCTC TGGCAGCCCA GTGCGAAGC CAGAGATGCG GACTCTCTCT 3840  
TAGCTCACT TCTGCTCTGT AGTGTCTGAG ACTGCGCCCA AGTGAAGTCT TCTCTGCT 3900  
GAGTAACTC AGTGAATGA ACTTGGACAT CCGCCAGTGC TCTCATATGC TAACTGAAT 3960  
CTGCTCTCTG AATTTGTTGG GGGGTGGATA GGGTGGGCTC TCACTCTACT TTTTGTCACT 4020  
25 TCACTCTGAA ATGCGAATAT ATGTAAATAA ATATATCAGC AAGC

Seq ID NO: 10 Protein sequence  
Protein Accession #: NP\_068830.1

30 1 11 21 31 41 51  
MRYTIDIFS GTWIDRLCY NDTFESLVLE GNVVSQLNIP DTFPNSKKY EHRHITMWW 60  
WRYRYKQKV LYRTETLIDA GCLSLNLRFP MDSRSCPLSF SSFVPEHME TYHVEPHLE 120  
THRGSHWLF QTTGQVSHR RL TLTTPVGD FWHHTFVPR SRFVQVQAF NVSPFVFM 180  
35 LSHVFWIKT SBAAQRLLE ILSVLDITLL TPTFRHIFPR VSYTALDFY IAIQFVFCF 240  
ALLFAVLIN ELYNQTKJHA SPKLRRHPRN SRBAHTRAR SRACARQHE AFVQIVTTE 300  
GSDDBERPC SAQQPFSPFG PEGFRLSCBK LACCEWCKRF KYPYCNVDFC BGSNQQORL 360  
CHVRLDMLY SRVFPVTFF FNNVLNLWC LML

Seq ID NO: 11 DNA sequence  
Nucleic Acid Accession #: NM\_001076.1  
Coding sequence: 22..1614

45 1 11 21 31 41 51  
TTGCGCAGGA GTAGACAGAG GATCTCTCTG AATATGAGCT CAGTCTTCT CAGATAGAG 60  
CTAGCTGTT ACTTTAGCTT TGGAGACTGT GGAAGAGTGC TAGTGTGGCC CACAGATAC 120  
AGCATTTGGA TAAATATGAA GACAACTCTG GAAGAGCTTG TTAGCAGGCG TCAATGAGTG 180  
ACTCTCTGGA CATCTTGGCG TTCTACTCTT GTCTATGCCA GTAATCATC TCGATTAA 240  
50 TTAGAGTTT ATCTACATC TTATACAAAA AAGGATTGGG AAGATCTCT TCTGAAAT 300  
CTCAAGAGT GAGATATAG TTTTGAJAA AAGCAATTT GTCTCATTT TTGCAATTA 360  
CAACAATTT GTTGGGAATA TTATGACTAC AGTAAACAGC TCTGTAAAGA TCGAGTTCT 420  
AATAAGAAAC TTATGATGAA ACTACAGAGC TCAAGATTTC ATGTCTATCT GCGAGTATCC 480  
TTATCACTCT GTGGGAGCTC ACTGCGTAAA CTATTAAACA TACCTTTCT GTGACATCT 540  
55 GATCTCTGT TCTGCTACAC ATTGAGAGG AAGGTPGAGG GATTTCTGCT CTGCTCTCT 600  
TATCTACTGT TGTATGCTC AGAATTAAGT GATCAAGAAG TTTTCAATGA GAGGATAAAA 660  
AATAGATAC CATATCTTTA TTTTGACTTT TGTCTCAAAA TTATGATCT GAGAGATGG 720  
GACCACTTTT AATGAGAGT CTAGAGAGGA CCGACTACAT TATTTGAGAC AATGGGAGAA 780  
GCTAAAGTT GGTCTATGCT TCTGCTGCTG TCTGCTGCTG TCTGCTGCTG CTGCTCTCT 840  
AATCTGATT TTTTGTGAGG ACTTCACTGT AAACGACCCA AACCCCTGCC TAAAGAAAG 900  
GAGAGCTTTG TCGAGAGCTC TGGAGAAATAT GGTATTTGTG TGTTTCTCT GGGTGTGAT 960  
TTCAGTACGA TGTGAGAGGA AAGGCGCAGC ATGATGTGCT OAGGCTTGC CAGATGCCA 1020  
CAGAGCTTTC TAGGAGAGCT GAGGCGCAGC AAGCGCAATA CATGATGTC CATGATGTC 1080  
CTGTACAGT GGTATACCCA GAATGACCTT CTGTGCTATC CGAAGACCAA AGCTTTTATA 1140  
65 ACTCACTGTC GAGCAATGCG CATCTATGAG CGGATCTACC ATGGATGCC TATGTTGGCG 1200  
ATTCCCTTGT TGTGATGACA ACATGATTAAC ATTTGCTACA TGAAGACCAA GCGACGAGC 1260  
CTGCTGCTGT ACTGACGAGC CATCTCAAGT AGTACTTATC TATCATGTC AAGATTTTGA 1320  
ATTATGAGC CTCTCTTAA AGAGAGATCT ATGAAATTTA GAGAACTACA TCAAGACCAA 1380  
CCATCGAGC CCTGTAGTGC AGCAGCTCTC TGGATTGAGT TTGTCTATGG CCAAGAGAAA 1440  
GCCAAGACCC TCGAGTCTCC ACTCTACAAC CTGACTATCA TCGCATACCA CTCTTTGGAT 1500  
GTUATAGAT TCTGCTGCTG CTGCTGCTCA ACTGAGAT TATCATGTC AAGATTTTGA 1560  
CTGTCTTGT TCGAGAGCTC TCGCAAGACA GGAAGAGAGA AGAAGAGAGA TTAATCTAT 1620  
CAGAGACTGT AAGTGAAGAT ACTGAAGAGT GSGACTCTCT CTTTATTCTA CACTGAGAG 1680  
TTTTAAATGG AGGATTTCTT TTTTCTGTGT ACAGAAACAT TTTTCAACAT TTACTTTTAT 1740  
75 AAGACAAAT TATTATTCCA GGGATTAAAT ACTACTTTTA GTTGAATTA TCTCATCTA 1800  
ATGATTTTGA ATGCTGAGC AATGCAATGTC GGGAGAGAT AGCTCTTGA GATATACCA 1860  
ATGTTAAAGT ACAGATTACT GAGTGAAGCA CCGACATGTC CAGCTGTGTA TACATATGA 1920  
CCTAAACCCCT CTGTGTGACAC ATGTACCTTA AACTTTAAG TATATTTTAA AAAAAGCAAA 1980  
AAAAAATAAT ACCACTCTT TTTTTTAAAC CAGGAGAGAA AATGTGAACA TGGAAACAC 2040  
80 TCTGATATT GGAATCGAAA ATAGATGTC ATCCAGCCCA TAAJAAAAA

Seq ID NO: 12 Protein sequence  
Protein Accession #: NP\_001067.1



WO 02/098358

PCT/US02/17594

```

1      11      21      31      41      51
|      |      |      |      |
1  MSKKTSVPL LQLSCLYPS GSCCKVLWVP TEYSIMIMMK YLLEELVQSG IIRVTLTSSA 60
2  STLVASKSS AIKLEIVPTF LTKHELEDSL KILDERWIVG VSKRTKNSYF SOLQELWMEY 120
3  YDSEKLCMD AVLNKLMKMK LQBSKFDVIL ADALNPCEGL LAELFNIPFL YSLRFSVQIT 180
4  FEICGGGLPL FESVYVPMHS ELSGQIIPME RIIMHIMLPL PDPWFQITDL KMDQFVSEV 240
5  LGRVTLFTT NGKASGAPL PLEWDFPAP LACKPAKPL KMSSEFQSG 300
6  GSHVIVFSL GSHISMSSE SANHISAIA QIPQKVLARF DCKPNTLGS KTLVYKMLPO 360
7  NDLLGHPKTX AFITHQOTNG IYBAIYHGIPI WGVPLFADQ HDHLMKMKX GVALSVDIRT 420
8  MSSLDLNLH KSVINDPYTK SHVKKLSREI HDQPKFELDR AVPMIEFPMVR HXGAKHLEVA 480
9  AINLVLIWAL SLEVIAPFLA CVATVIFILT KFCLEPCTREU AKTKKKKKED

```

Seq ID NO: 13 DNA sequence  
Nucleic Acid Accession #: NM\_014109.1  
Coding sequence: 651..1739

```

1      11      21      31      41      51
|      |      |      |      |
20 CTCTCATCA TCCTTTGGA AAGCTTACGG TATATACAT AGACATCTCT GTCTCTTTTG 60
21 GAGTAGATAC TACATCCCTT GAAGAACAAT TCCCAAGAGT GATTCTGGA GCTAAGAGAA 120
22 CAGCAACAAG TATAGTGAT GTTCTTCATA TCCACGTGTG GTGGGAAATA GTTGGACCGA 180
23 CACTTAAAGC CACATTATAC ACATATTATC AGAATATTC TCACTTTCCT CGAGTTTATC 240
24 TACTTGCAC TTCTGCAAA CCCCATTCCG CTTCGCGAG AGAGVIGCAA GAATCTTTTA 300
25 TCCCTGATTA TGGAGAGATT TTAAATGTC AGTTATCCGA TAAAGAGAGA CGACAAATAT 360
26 TTTTGTGAGA TTTAATTCTA AAACAACCTG CTAAGCCCTC TATATCAAAA AGCAAAGCAG 420
27 TTTTGAGGCG TTTGAGGATA CTCCCAAGTAG CACCAACCACT TGGACCAAGA TCATCGACAG 480
28 CAGAGAAGAT GAAACGCTGA GAAGAACGAG AAGAGAGATAC ATTTAGAGAA CTCAGGATTT 540
29 TCTTAGAGAA TGTATACACT AGCTCTCCFA TTGACACACG ACTCCAGATG TTTACTAGAG 600
30 CTCTTGACCC TGATGAGGTT CTGATATATG TCACTTATAT AAAGCAACCA ATGAGACCTT 660
31 CATCTGTAAT CAGTAAATTT GATCTACACA AGTATCTGAC TGTGAAGAGC TATTTGAGAG 720
32 ATATGTGATC AATCTGATAT AATGCTTTAG AATCAACCTC AGATAGAGAT CTTGAGAGAT 780
33 GTCTTATAG GCTAAGAGCT CTATCTCTTA GAGATACCTG CATTCACATA ATTAAGAGAG 840
34 AACTGTATCA AACTCTTGAC CAGCTCTTAG AGGAATTTCA GGAATCTAGA AGCAAAGCAG 900
35 GTGTAGCTTC CTCGAAATAT GCCCGTCTTT ACTACAATST GATGCGAAAG CAJAATTCOA 960
36 CTCTTGTGTC TGAATAAAGA TCAGACACAG ACGAGATATGA AAGCTTAAGG ACACCGAGTA 1020
37 CTCCTGAGCC TGTGAGCTCT CAGCTCTCAT TCTGATCTAG AATTCGCGAA AATCAAACT 1080
38 GGTACTAGGG CACCAATAAA AAGCGAAGGA AGATTTTACA GCGAAAGGAT GATAGCCAG 1140
39 ATGCTATAGA TCACAAAATT GAGAGTGATA CAGAGGAAAC TCAAGACACA AGTGTAGATC 1200
40 ATATAGAGAC CGGAACAACA GAGAGTCTTT CCGTGGAGAA AATGAAAAAA CAGCAAAATG 1260
41 GCTCTGAGAG CAGTACAGAA TGTGAGAAAT ATTCAAATAC TTGTATATGA GAGATATGAC 1320
42 TGTAGCTCTC TAGAGAACTT AACCAATATA AACCAATATG ACGCAATTT CCTGTATATG 1380
43 GAGATGCTTC TAGCTCTCAG ATATATCATA TTTCTGATGA AATGAAAGGA AAGAAATTT 1440
44 GTGTTCTGCG AATGACTGCA GCTAGAATTT CCCAGGTAGA ACAGCAACAG CTCTACACTG 1500
45 TTGAAAGGCC TTTCGAATT CTCTTTCGAC CTACACCTTC ACTGTGTGTG GATCATGAGC 1560
46 GATTAAAGAA TCTTTTGAAG ACTGTTGTGA AAGAAAGTCA AATCTACAC ATATTTCAGT 1620
47 TGGAAATTTT GTATCAGATA ATCAACCAAT GTATTATGCG GCATGCCAAG GACCATGATA 1680
48 AAGCATCATT TATTCAGAAA ATGAGGCAAG AGGTAGAAAA CTTCACTTGT TCCAGATAGT 1740
49 GATGTCAIAG TATCAGAGAT TCTTTATATT CAGTTCTCAT TTAACTATT TTTCAGAT 1800
50 CCGCTATANT GATGTAGATT GAAACCTGTC ATCTTTAGAG AAGAGATATA AATGATAAA 1860
51 TAAAGTATT TAAACTTCTC TGAATTATTAT GTACATATTA AGATAAATGT CATGTGTAG 1920
52 ATAACTGATA AATA

```

Seq ID NO: 14 Protein sequence  
Protein Accession #: NP\_054828.1

```

1      11      21      31      41      51
|      |      |      |      |
60 MDLSSVISKI DLHKVLTVID YLEIDILICS NALEYNPRDR PDRLEIRHRA CALRDYAVAI 60
61 IKELDORFDE GLCEIQESR KRGCSSESY APSYHVMKPK QUSTLVGDKR SDFRQEKELK 120
62 TPSTPVACAT PAQLKRIKIR KHWLYLTIK IERNKIQAND DQMAIDHKK ISDTSTGDT 180
63 SVHMDTQNT GATLTGAGL QGASSEKSLK LKHWSTQCTI ENHEDERTCT TACTALAKKI 240
64 ACNDASSSO IIRISIDNKG KENCVLNTR ARRSVROOD LITVSKALAI LSQPTSLVV 300
65 DIERLNLKLL TVVKKSQNYN IPQLEHLYAV ISQCIYRHK DHDCTSLQIK MEQVEVFNSC 360
66 SR

```

Seq ID NO: 15 DNA sequence  
Nucleic Acid Accession #: AK001536

```

1      11      21      31      41      51
|      |      |      |      |
70 TATATGTGAC CTTTATAAAA AATGACTGTG AAGCATCTTC CCAAGACATA GCTCAGCTTC 60
71 CAGAACTCTC TTTCTGATA GTTGAGACCC CTTCTACACA CAGATGGTA GCAACAATTC 120
72 ATAGTTCGAA GTTGACGAAA TTGCACGAG ATCAATATGA AATCTCTATC AATCACTTCA 180
73 CTCGAAACCT TATCTCAAGT TTATCTCCA AACCAAACT TCTCTTAGAC ATCAATCTCA 240
74 AAGAGCTCAG AATCCAGATT TGGTTTCAGA ATCAAGAGAC TAGCATGGA TTCAAGAAA 300
75 CACGAGAAC CTGACTTTAGA TTTAAOCCAC ACCCATGGAC AAGATTAACT TGTGTGAG 360
76 TTTCAAAATA GMAAGACAG ATGCTGTGTG ACCGCTATTA GCACCTTTCA ATTCACACA 420
77 ATGCTGATC GATTTGAGA AAACCTATAC CTGCGGATG ATTCGCGGA ACACATCTC 480
78 GAGAAATGTG GTGCTTCAGA GTCAGAGTTC CAAATTGTGT TCCAAATATC AAGATCTAGA 540
79 TTTCACTCTC AGAGAAAGAG AGAACCTGT ATGCTCTTAG AATGAGAGA CCGAGAGAGA 600
80 CCAAGGGGGA GTTCTTCTAG GAGCTCTCAG GTACAGAGA TACCAAGAT GCACACAGC 660
81 TCCATGACC TCTCATTTCT CAGAGGACAG AACATYGGTA ATACACATA ATTCATATTA 720

```

WO 02/098358

PCT/US02/17594

TTTGATATA TCAATTTGGG CCCCAAATCT CTCTCAGACT CTCTCTGGGA GTCTATTCTT 780  
CTTCCAAAG TCGAAGCTGA GCTTCTGGA ATGTGTAAAG AACTTGCCG GGTGTGTGGG 840  
CTATGCGCTG TATCCCGAGC ACTTTAGGAG GCTGAGGCTG GAGATTTGCT TGAGCCTAGG 900  
5 AGTTTGAJAC ACTCATGGAG AAGTATGTAA GACCCCTGCT CTATTCTTAA AACAANAATA 960  
AGTAAGAAG ACTGTAGAG GCGAGACAGA GTACAGGCT GACACACTA CCGCTGTAC 1020  
ACAGCCGTGA TCCGAGCTTC AGCAGACCTT GAGACAATBA AACAAGACTT AGTAATAATC 1080  
ATTTTTCAT CATTCAGTA ATTAATGAT TTGACAAAJA TCAATTGAGC TCAAAACCTT 1140  
AAGTGTACT TTCTTGCTCT ATGAGTGTGT CATCTCTTTA TTCTTTAGT TTCAATAATA 1200  
ACTTCTTTTT ACTTAATAA ACTTATATTT TGAGTAGAGT TGAGTAGAGT ACTGTATCTT 1260  
10 AAGAAATCT TACACACAGA CTATTAATTT ACAAAAGGAA AATGCAATAT TTGTGACGTG 1320  
AGACATATGG CCAATCCACG CTATCCCGAG TGCTGGAAGT TCACTTGACC AGTAATGTGA 1380  
CAAAACAATG TGAGATGACT GCTGATATGA GACACTTAAG AGGGCAGTGT CTCTCTGTGA 1440  
TTCTGACAG AAAAAATGG TTAGCTGTAC CTGAAGACTA TTAATGACTA ATGTGATGAC 1500  
15 AATACAAAT GTATTTGAGA GTCTGGAJAG TACATCCCAT GTACTTTCTA ACATATGATC 1560  
GACCTTAGCT ACTCAGGAGG CTAGAGTGGA ATAAATGTTT GAGGCACAGA GTTCCAGATC 1620  
AGCCTGGGCA ACATCATGCG ACCCATCTCT TAAAAAACATC TTTTAAANAA TGAGCCAGGT 1680  
GTGTATGACT GACCCGCGAG TCTCAGCTAC CTGAGAGGCT GAGGCAGAG GATGGTTTCA 1740  
ACTAGAGAG TCGAGCTGCG TTGTAGCTA GTGTGTCTA CTGACATCGA GCGTGAGTGA 1800  
20 ACAGGACAT TCTCTTTTCC AAACAACACG AAGAAACAGA AACAAAACAA AACAAAAAAT 1860  
AGATAGATA GTGACATAAA AATGTGAGAA AAGTAGGCT GACTCAGGAA ATGCTTAGAA 1920  
AGTACGCCA TACTCAAGG ATATTGTAGA TTGTGTCTGA GACCCACAGA ATAAAGCAGA 1980  
TTATGCGCA AAGTAGCTGA CAGAAATGTT TTTGTCTCT TTGTATATAT AAGTATAT 2040  
GGCTGGTGTG TAGTGCTCAT GCTTAATATC CAGTACTTT AGGACAGCGA GCGGGGAGGG 2100  
CTACTTGAGC CAGGAGATTG TGAGATCAAC CTGGCCATAG AGGAGATGCC TGTCTCTATT 2160  
25 TAAAAAAGA AGTATGTTT ACATACACT ATAGTCTATT TAAGTGTGA ATGAGGTTA 2220  
TGTCTCTAA TTTAAAGCT TCGAGCTGG CTGGCTTGGG TGCTGTACG CTGTATCTCC 2280  
ATCATCTTGG GAGGCAGAGA CAGTGTGAT ACTTAAATTC AGGATCTCAA GACCCAGCTG 2340  
GACACATGG CAACAACAT CTTTAAAAA AAGAAAGAAA AAGAAAAC AGAAGAAAAA 2400  
30 AAGAAAAA CTACTTGCT CCTTACTTG AGTCTAAT ATTAAAAA

Seq ID NO: 16 DNA sequence  
Nucleic Acid Accession #: CAT cluster

1 11 21 31 41 51  
35 | | | | |  
CTTTTTTTT TTTTTTTT TAGTAGAGAC AGGGTTTAC CATGTAGCC AGGATGCTT 60  
CGATCTCTG ACTCATGAT CTCTCTGCTT TGCGCTCCCA AGTGTCTGG ATTACAGCG 120  
TGAGCCACTG CACCGAGCCG AGAGTTTTT TTACAGAGT TCTCTCAGC AATCTGACTA 180  
40 TCGAGATAA GGTTCATCT AGACATACCA CAGGCTHTA CTCTATATCT CTGTGATA 240  
CAGAGTTTC CTGAGCTCTG TGATAGCTCA CTGCTTCCG TCTATATCA GTGTTTTAC 300  
CTGATGTGAT TTGTAGTGTG GTTCTGTGCT TTCTGTAGCG AGAAAAAAC TTTCATTTT 360  
TTTTTGCTTA CATGACATA AATGTAGCG CAAATCTTA TATTAJACT TTTATTCTTA 420  
TATAGCTAA TTCTGTGTT TCTGTGCTA ACCAACCA GAGTATAGG AATGATACC 480  
45 TCAAAACTG ACTAATTAG AGATCAATA ATGAGGCTGT TTAATTCTA TATCTCTCT 540  
TTCATGATT AATAGAAAA TTTT

Seq ID NO: 17 DNA sequence  
Nucleic Acid Accession #: CAT cluster

1 11 21 31 41 51  
50 | | | | |  
GGACGAGAA GAGCCACAT CCCCATTAT AGAAGGCTA ACTAAATTC ATGATCAGAC 60  
ACTAATATT GTTCTCTAA TTAGCTCTCT AGTCTCTGAC ATCATCTGCT TATATTATAC 120  
AACAATAAT ACACATAGA GCAATATGA TCCACAGAA GTGTAAACA TTGTAAACTA 180  
55 TCGACAGAT GTTCTCTGCT TCAATATCT CTAGCTCTCT CTAGCTCTCT TATATATAT 240  
ACAGCAATCT AACAACCCG TATTAACCTT TAAACCATTA GGGCCCAAT CATACTGAG 300  
CTACGAATAT ACTGACTATG AAGACCTATG CTTTGATTA TATATATCT CAACAAACIA 360  
CTAAACACT GGTGACATAC GACTGCTAGA AATGTATAC CAGTGTCTT TGCCATAGA 420  
60 ACTTCAAC CAGTATATA ACTGATCTCT AGATCTCTCT CTGACATGAG GATGCGCTC 480  
TACAGACTT TAACTGATG CCAATCCCGG CCGACTAAT CCGACAGAT ACTACACCG 540  
ACCGAGTTA TTCTATGCG AATGTCTGA TTGTGTGCT TACATAGCT TTTTGCATT 600  
GTCTAGAAAT GGCTCCCTAA AATATTGAG TACTGTCTG

Seq ID NO: 18 DNA sequence  
Nucleic Acid Accession #: CAT cluster

1 11 21 31 41 51  
65 | | | | |  
GTGTACATCA GAGCAAAAT ACAGAGTAT TATTCATTT TCCCGACTAG AGGACACAC 60  
70 TGTCTTGA GACACAATG AATCATGAT TGTGAGAGT TGCTTTGGA GATGATCAA 120  
TGAATCTCT TTAGTGCTT GGAATGTAT ACAGGCTCT ATCACTCTG GACACATCA 180  
GCTGTGAGC TTCTGTGCT TTGAACCTA ACTGAGCTCT CTGAGCTCT TATATATAT 240  
TAGCAAGGT TTCTGTCTT TTCTTTGAG GGCATCATTA GAAATTCGG GGTATCTCT 300  
TTGATATCG ATCTGTGTA TACTCCCAT GTTACGAGAT GTAGCTCTT GGAATTTCT 360  
75 CAGACTGCG TTACCATATA AATATGTCTA CAGTGTATAT GCAATATATA ACTGGATCT 420  
CACTCAGCT AAGTATATA TTATATTAAT CAGTGTCTCT ACTCTATAT TATAGAGAG 480  
GCTTACGAC AATCACTCT TTGCTCTCT CCACTGTGCT GACACAGAG CAGAGCTTT 540  
CGAAATATCC GGTGAGGAA ACTTCCGACT CAGAGTCTAG GACCGCGCG GCGCAGAGAC 600  
80 CACCTCTCTA GCGCGAGAC CBAANCGCT GAGCAGCTG ATGCGCGCA TGG

Seq ID NO: 19 DNA sequence  
Nucleic Acid Accession #: CAT cluster

WO 02/098358

PCT/US02/17594

1 11 21 31 41 51  
TAGTCCAGTN AATTACTTAA ATTTCGCTT TCCATATATC TGGTATTCOA TAGAAGAAAA 60  
TCTTTATTA TACTACTCTA CTATCAATCT CAGACAACGA TGGATATAGT TCCGATATGT 120  
5 CCAAAATTCG GAAACACCAT TAAAGGCAAT TCATATCTTA TTTTGGCAGT ACTCATTTCA 180  
TCATATCTTT ATTTCGATCG TAAGCGCAAGT ACTTCTAAGG AAGAGCTGTC ATATATATAC 240  
TTTAGTAAAG CATTCAAGTAD AGGCAATATAT CAAACCTCTA TCCCAATCTT TACACATCTG 300  
ACAGAGTAAG AGGAGGATGT AGCAACATACA TTTTTCGACA TTTTCACTATTA AGGGCCATTA 360  
10 TCATTTAGG GGGCTTAGG GGGCAATATAT ATATATATAT ATTTTGGAC A

Seq ID NO: 20 DNA sequence  
Nucleic Acid Accession #: U93672  
coding sequence: 351..3761

15 1 11 21 31 41 51  
GCCGCGGCCG TGGGCTGAGC TGGAGGCCAGG CCGGCGGACGC CGGAGAGGAGC TGGGCGGCTG 60  
GGGTCTATGG ATGCGCGCGC AGCGCGCGCG GAGCGCGGGA GCGCGCGCGC GACGACGCTG 120  
20 GGAGCGCGCG TCGCGCGCGA CTGACGCGCG GAGCGCGCTG GAGGAAAGAG GCTCGCGCGC 180  
CGCGCGCGCG CGCGCGCGCG TGCGCTTCTC GTTGGGATTA TCTTCTGCTC CGCGCGCTCT 240  
CTCGCGCGCG CGCGCGCGGA GCGCGCGCTG GCGTCTCAAG CCGCGCGCTC CTGGCGCGCG 300  
GGTCCGCTCG CTACCTGGCG CTGTATGCTG GAGTACCTCT GCGCGGACGC ACAGGAAAGT 360  
TCAACATCGA GAGGTGCTG GAGCGCGCTG CGCGCGCTC GTTCTCGCGC TCGCAACAGC 420  
AGCAACAGCA GCGCGCGCG CTGCGGAGCC GGGAGCGGGA GATCGAGAGG AGGCTCCAGT 480  
25 CCGAGAGCTC CGAGCTCGCG AGAGCTTGTCT CCGATGAGAT TCGGATCGAG CCGTGGAGCC 540  
TGGCTTTTGA TCGCGCTCGA AAGATCTTGG GCGTAAAGAC CCGAGACTGT CCGTTAAGCG 600  
TCTTTGGTGC TCGAGGGGTG GAGTGTGATC CCGACACAGA CAGCGGAGCG CGAGTGAGCT 660  
AAGCTCAGTT CCGATCTAAT GAGGGAGGCC TTGTGAGCGC CTGCGCGTAT GAGACTCTAC 720  
30 ATCTTGGGAA TTTCATCTCTA AAGAGCGGCT CCGTCACTGA CTGCTGAGAG TGTTCAGAG 780  
AAGGGTAACT ACTTGGCAT CTGCGCTTCC AGACTAAGTG GCTCATATGT GCGACCGAAG 840  
GAGGTAAATC ACATATGTC AATGTGGAGT CCGTCACTGA CTGAGCGTAC CTCATTAAGT 900  
GAGATTAAGC CATCGAAGCT TCACTTAAAT CTACACCGAG AGCTGTGCTC CATATAGTGT 960  
ATATCTGCTC GCGTCTGCTA ACTCTCTTGA ATCTCTTATC ATCTCTAGAA CTCTCTAGAT 1020  
GGGACTAAAT TCGTAAGAGG CTGAGCTACA CTGAGACTTA CCGAGCGGCT ACTCACTCTG 1080  
35 TGGCTTGGCA TCACTAGAGA AAGACGTTTA TTTCAGTGA TTCTGATGAT ACATTGACGA 1140  
TATGAGATGT GAGGTCTCCT ACTGAAAGCT CTGAGAGCGT CACTCTCTAC GGAAGACAGT 1200  
TAAAGATGTG GAGAGAGTGA CAGCGCTGGA AGCTCTATCT CAGGTGCGAG TCGAAGAGTA 1260  
CAGAGTCGCG GBAAGCTTCT ACTATTTGTT GCGGAGGCTC ATCATATGAT ACTCTGGGA 1320  
GAGAACCTGT CTGACAGTGT ATGCAATGGA AAGAGCGGCG AGTGTCTGAA ATGGAATATT 1380  
40 CATATTCGCA CTCTCTACCA GCTGTGAAA GCGCATATCC AATGATTTT CAGGAGCGGT 1440  
ATCTCTGTGT TGTCTCTCTG GAGGAGGACT TAGTCTGAG AGACTCTGGA CAGAAATGAG 1500  
ACCTCATATC TGGAGTATCT ACCTCTTCTA GTACAACGA GTCTCTGTTT ACATTTGTT 1560  
ATATTTTCTC TGAATGTCTT GTGAGCTCTA TCTCTGCA CT TATCTGTGT GAGGTAGAG 1620  
AGAAAGCTCA AGCTTACAGC AAGAGGAGAT GCGCCATCTA TGGTGTAT TGGGGCTTGG 1680  
45 GTGCTCAAG TTAACGAGGA ATAAATATTA CAGCGCATGC TGTGSGTCA ATTAATATT 1740  
GGGTCTCTTC TCGCAATATC CTACAGTAC TGTAAAGTAT AAGAGATCT AAGATATTG 1800  
AAGAGTCAG AATTAAGAT GACAGACAGA CACCGAGAT TGTGATGAA GATCTCATTTG 1860  
CCATTCAGAT CATCTCTGCG TCGCAGAGGA CGAGATGCTT GTGATAGGCC GATCTCATTTG 1920  
CTCATGTGAC CATTTCATGA TCTCAGGAGC AGGAGATGTT TACAGAGAGT ACTCCGATGT 1980  
50 TCGAGTCTCG ACCTGTATAT GAAATTAATG ATGTGAGAGC CGCGGAGGCT CAGCAGCCAC 2040  
CGCTCTTGTG CACTCCGCTG GCGAGCTCTA CCGTCTGAGC CATCCCCCT CAGTCTCATCT 2100  
CGTCTACAG CAGCAGCTCA TCGAGCGGCG TTGAGATGAA TGTACCGTGT TTAAGATTTA 2160  
AAAATCAAC ACTTAAGAG TCTCCCGGCT ATCAAGAGGA GGTATGATCT CAGTTGTGTT 2220  
55 TCGAGTCTCG AGATATCTCT CAGGAGATCT CCGAGCTCTA CAGTCTGAGT AAGAGAGTCT 2280  
TGGTGTCTCT GCGACATCTC AATGAGATCT CAGTGTGTTA CTACTCTGAG AAGAGAGTCT 2340  
TCTCAACTCT CAGCAGCTAT GAGCTATAGC GCGCATATGA TCGTTATGCG AAGAGAGCTA 2400  
GGTGGCGCG CAAATCTGGA CAGCACTCTAG GAGCGGCGCT GGTGATATT ACCGAGAGGA 2460  
TGTGCTGCTC AGAGATGATG TCGACTCTCG CAGCTCTGCG AAGATATCTA AGGATATTTA 2520  
60 GGTCTCAAC TGAATTAAG CCGATATTAG ATGTGAAAGA CAATCTCTCT ACAGAGATCT 2580  
GGAGTCAAG TGTGACAGC ACTGACAAAG AGTCCCGGGA AGCCATTCT CCGTCTCATT 2640  
TCTGTGAGAC TTTCAGAGG AAGGAGAGCT CCGCCCGCTC CCGTGGTGTG TGGGTGGGAA 2700  
CTAGAGTGG AGGATGATCT GTCTCTGCTC CAGTCTGCTC CAGAGAGTCT CAGAGAGTCT 2760  
70 TCTCTGAGC AGTGTATGTT TCTCGAGAGG GTACTGATAT GAGGTAAAG GTGTGCTAT 2820  
TGGAGATGCG ATTTCTGATG TCGCGGCGCT GCTTATGCG ACCGTGATAC GAGCGCTGGA 2880  
CGCTGACAA CCGTCTCTAA GAAAGAGAG ATGAGAGAGA ATGGAAGAG GCGGAGACCTG 2940  
65 TCTCAGTCTC GGTCTCTCTC TCTCAGGAAA TTATGTGAAA CCGTGTGAG GTATATATTT 3000  
CTCAAGATCA AGCAAGATCT CAGTCACTG CAGCTCACTA AATCAGAG CTGACCTACA 3060  
TCACTGAGAC GTCTCTCTG TCGCTGGAGG ACACTTGTCT CCGTGAATC AGGCTGTGCG 3120  
TCTGCTGCT CTGCGGACAG GCGACATTTA TCACTTTGAG TTTCGCGAGC TTGAGGCTCT 3180  
75 TCTGAGTCTC TCTGAGTCTA ACATGCGAGT CAGTCTGCTA TCTGCTGCTC TCTGCTGCT 3240  
CTCAAGTCTG CAGAGCTTTA TACTCTTTCT CAGTCTCTA AATCAGAG CTGACCTACA 3300  
GTCAAGAGAC GTCTGAAAGC CTGACGAGA TCGTTGGTGA GCTCTCTCAG CCGTGAAGAA 3360  
CAGCAGAGC ACCAAGAGGA GGGTCTCTCA AAGCTCTATT TGGAGTGTGT GCACATCTCT 3420  
TTGATATAGA AGACATTTT GAGAGAGCTC CCGTGGAGAA GCGCTCAAGG AGCTCTGAC 3480  
80 ACAGCATCTC GGTCTCTCTC GGGTCTGAG GAGTCTGAG AGCTCTGAG GAGTCTGAG 3540  
GAGATCTGCG CTGAGCGAGC CTGCGGCTCT CAGAGAGAG ACAGAGCTC AGCTGATCTG 3600  
AAGAGAGAGC TGAAGCATCT ATGTCCAGTGT CAGAGTCTGT TTCCAAACAT GCTCATAGGA 3660  
TCTGATGAAA ATACAGAGAT AAGAGAGTGT ACCAGTTCTG ACAGATAGCA CTGAGTAAGT 3720  
CTACTCTGAA CCGAGAGAGA AAGAGATTTT CCGTTTGAAG CAGTCTGAT TATTGGGAA 3780  
ACATATCAT AAGAGATCT ACAGCTCTGA CAGCGTCTT CCGACACAAA TCACTACCT

Seq ID NO: 21 Protein sequence  
Nucleic Acid Accession #: AAB84756



WO 02/098358

PCT/US02/17594

5  
10  
15  
20  
25  
30  
35  
40  
45  
50  
55  
60

AAAAA AAAAGCGAAAT AAAAGAAAAA GATATAAAT CATTTCTTAT TTGCACCTAC 780  
TTGACGTGGAC ACTGAATTGT GAAGSTGGAG GATTITGTTT TTTTCTTTTA AGATCTGGGC 840  
ATCTTTTGA TCTACCTCTT AGATATTAAG AGACAGACTG TGAGCTTAGC AGGGCAGACT 900  
TTGTCACCG TGTGCTCTCT TCTGCAAGAG ACTTTGAGG TGTGAGAGCG CTCTTGGCT 960  
GGTGTCTCCG GCGAGTTCTG TTTCTTGGAG GTTCCGACAG GTGTGAGGCT AGCTCCGAGC 1020  
ACTACCGCAT CATACAGCGC TGTTTAACTC TCTTGAGCAA GAGAGAGGGG GGGGGGTAA 1080  
GGGAGTAGAG TGGAGAGATTC AGCCAGCTCT AAGAGATGAA GTGCAATTAG GGTCTGGAGAG 1140  
GGTCACTCT TGGAGAGCGGT CAGAGAGCTA CCGAGAGACT TTCTGAGAGC TGTCTCAGAG 1200  
GCTGACGAA GGTGATCGAG AATCCGCTCA CAGACATCCA GAGTGTGCTA GGTGATGACT 1260  
TCCCGCGCCG AGTTTGTGCTG TCTGTGAGCA CAGACAGGAG CAGCAGCAGC AGCAGCAGCA 1320  
GCGACGAGAG CAGCAGCGAG AGCGAGNAGA GACTACGCCG AGGACGAGAG CAGACGAGCA 1380  
GGTGTAGATC GGTCTCTCCC AGGCCATCTG TAGAGSCCCC ACAGGCTACC TGTGTCTGGA 1440  
TAGAGAGAG CAGACTCTGC AGCTGAGAGC GGTCTCTGAG TCTCTCTGAG AGGAGAGTGT 1500  
GCTCCAGAG CCTGAGAGCGG CGTGTGCGCC CAGCAGAGGG CTGCGCGAGC AGCTCCGAGC 1560  
ACCTCGGAGG GAGAGATGACT CAGCTGCCCC ATCCAGCTGT TCCCTGTGCT GGGCCACTTT 1620  
CCCTTGATT AAGCAGCTGCT CCGCTGACCT TAAGAGACTC CTGAGCTGAG CCGGACCAT 1680  
GCACTCTCT CAGCAGAGAG AGCGAGAGAG AGTATCCGAA GCGCGAGGCA GCGTGGAGCT 1740  
GAGGAGAGCG TCGGGGCTCT CCACTTCTCT CAGCAGCAAT TACTATGCGG GACCTTCGAC 1800  
CATTTCTGAG AACCGCAAGG AGTTGTGTAA GCGCATGTGT GTGTCCATGG GCTCGGTGT 1860  
GAGAGGCTGT GAGCATCTGA GTTCAGGGGA ACAGCTTCCG GGGGATTTGA TGTATGCCCT 1920  
ATTTTGTGA GTTCCGCTCG CTGTGTCTCT CACTCTCTGT GCGCATGTG CCGAGAGGAA 1980  
AGGTTCTCTG CTAGAGGACA GCGCAGGCAA GAGCATGTGA GATACTGCTG AGTATTCCCC 2040  
TTTCAAGGGA GGTTACACAA AAGGGCTAGA AGCGAGAGAG CTAGCGTGTCT CTGAGCAGGC 2100  
TGGAGCGAG AGCTACCGAG CATCTGAACT TCGGCTACTG TACACTTCTT CACTGTCTCT 2160  
AGCTCTGGAG GAGTCAGCTG CTTACCGAGG AGCTCTGCTG TACACTTCTT CACTGTCTCT 2220  
GGCGAGAGCG CGCGCCCTCT CGCGCGCTCC CCATCCCGAG GCTCGAGTCA AGCTGTAGAA 2280  
CGCGCTGAG TACGAGCGAG CCGTGGCGCG CAGTGGCGCTG AGTGGGACTC 2340  
GGCGAGCGTG CATGGCGGAG GTGCGAGGAG ACCGCGTACT GGGTCACTCT CAGCGCGGCG 2400  
TCTCTGATC TGGACATGAG GTTCTACAG CAGTGTGCTG CAGTGTGCTG GACCTGTGCT 2460  
TGTGTGTGAG GTGTGTGTGAG GCGCGCGGCG CGCGCGCGCG GCGCGCGGCG GCGCGCGGCG 2520  
CGCGCGCGCG GAGCGGAGAG CTGAGGCCCC CTAGCGCTAC ACTCGGCCCC CTGAGGSGCT 2580  
GGCGGCGCGG GAAAGCGACT TCAGCGGAGC TGAATGTGCG TACCTGTGCG GAGTGTGAG 2640  
AGAGTGTGCG TATCTGTGCT CCGTGTGCTG AGTGTGCTG TGTGTGCTG GAGCGGAG 2700  
CTACTGCGA CTACTGCGAG ACATGCGTGT GAGCAGTCCG AGGAGACGTA TTTTGGCAT 2760  
TGACTATTAT TTTCCAGCCC AGAAGACCTG CCTCATCTGT GAGATGAGG CTCTCGGTGT 2820  
TGACTATGGA GCTCTCACAT GTGAGAGCTG CAGAGCTCTT TTCAAGAGAG CGCTGAGAG 2880  
GAAACAGAG TACCTGTGCG GAGCGAGAA TACTGTGACT ATCTCAAAAT TCCGAGGAA 2940  
AAATGTGCTA TGTGTGTGCT TGTGGAATCT TTATGAGACA GAGTGTGCTG GAGTGTGCT 3000  
GAGCTGTGAG AAATCTGTGA ATCTGAACTC ACAGGAGGAA GAGAGGCTCT CAGACACAC 3060  
40 TACGCCCATCT GAGAGGAGAA CCGAGAGCT GAGCGTGTCA CACATTAAGG GCTATGAATG 3120  
TAGCGCATCT TTTTCTGAGT TCTGTGAGAG CATTAAGCTA GGTGTADTGT TTGTCTGAG 3180  
CGACAGCAG CAGCTGTGCG CAGCTGTGCG AGCTCATGAT AGCTGTGAG 3240  
GAGACGCTT GTACAGGTGT TCAAGTGGCG CAGAGCGCTG CTGCGCTCTC GJACTTAA 3300  
CGTGAAGCAG CAGATGCGTG TCATTCAGTA CTCTGTGAGT GGGCTCATGT TGTTCGAT 3360  
45 GCGCTGAGCA TCTCTCAGCA ATGTCACTC CAGAGTGTCT TACTTGCCCG CTGATCTGCT 3420  
TTTCAAGAG TACCGCAGCG AGAGCTCGTG GTTGTGAGC CAGTGTCTG GATGTAGCA 3480  
CCTCTCAG GAGTGTGAT GTCTCGAATC CAGCCCGCAG GAACTCTGT GATTAAGGCG 3540  
ACTGTACTCT TTCAGCATTA TTTCCAGTGA GGTGATGATA AATCCAAAT TCTTTGATG 3600  
50 ACTTCAAGT AGCATCATCA AGGAGCTGGA TCGTATCAAT GGTGTAGAAA GAAATAATCC 3660  
CAGCTCTCTC TCGAGAGCTT CTGATGAGCT CAGCAGCTCT CTGATCTCTG TCGAGCTCT 3720  
TGGAGAGAG CTGATCACTG TCACCTTTGA CTCTGTAACT AGTCAACACA TGTGTAGCT 3780  
GQACTTTCCG GAAATGATG CAGAGATCAT CTCTGTGCA GTGCGCAGA TCTTTCTGTG 3840  
GAAATCTAG CCGACTTAT TCCAGCCGCA GTGAGGATAT GAGAACCTTA TTTTCCGACC 3900  
55 CAGCTCTCTC CAGCTCTCTC GATCTCTCTC AGCTGTGACT ACTCTCTG TCTCTCTCT 4020  
AGTGTCTG GGAATTTCCCT CTATTGATCT TCTCTTCTCT TCTCTTCTCT TCTCTTCTCT 4080  
TATTTGCGAG TGTCTTTTTT TCTCTTTTTT TCTCTTTTTT TCTCTTTTTT TCTCTTTTTT 4140  
AACTCTGCA TGGCACTCTC AGTACTGTCT TCGTATGTG GCTCTGACT CTGTGTGTGA 4200  
TGTGTGTA TGTCTGCTG TGTGTGTGCT TGTGTGTGCT GCGGAGCTC AGTGTGTCT 4260  
60 TGTGTAGAG ACTACTCTGT GCGGCGGACA CAAAGCTTGA CTATCTATCT GCGCGGAGAA 4320  
GTTTAGAGG CTAGATTAT CTGGGAAAT CAAAGCAA AACAGCAA CAAAGCAA 4380  
A

Seq ID NO: 25 Protein sequence  
Protein Accession #: NP\_000035.1

65  
70  
75  
80

1 11 21 31 41 51  
| | | | | |  
MEVQLGLGKVP YPKPSKTYK GATQNLQSV REV IQNPGR IPEHSAAPP GASLILQQQ 60  
000000000 000000000 SPRO00000 BDSGPQHRH GPTSLVWDE IQPSPQPSA 120  
LDECFERBCG YPKPSKTYK AGTQNLQSV FORDSDAHS TGLSLPTFP GLSSCSADLK 180  
DLEGAEDQ LQEQQDQAN SRSRSRHAR BASGATPSK DDTLDTSTI DNNMLTCA 240  
VBVSGVLGE ALHLGLPGSG IAGDCMYAPL LVVFPVAVPT PCAPLAECKS GLLDGGAGS 300  
TETZAEYSPF KGGTGTGLGE EELSCGSGSAA AGSSTLTLEP TSLSLYKSA LDEAAATYQR 360  
DTYFFPLALA GPPPPPPPHF FHAILLQHF LUTGSAMAAA PAQCKTGDLA SHLAGAAGP 420  
GGSPHSAJS SSMHFLPAB EQVYVPCG GCGGCGCGCG GCGGCGCGG GGGHGVAPY 480  
CTKTPPGDLA GQSDPTAPD VVYQGVNVR VPTSPCTCK SENQPMWGS SGPVGLMKLE 540  
TARDYHLDL YTFPQKTL ICGRHAGSK YGALCTGSKG VPKRAAEK QKYLCAHSND 600  
CTIDFKPURN CFSCKRKYT EAKTGLGRK LKXGLMLKLT EHGASSTTS PTBTGTQKT 660  
VBIITGTCG TITLPLPAB EQVYVPCG GCGGCGCGCG GCGGCGCGG GGGHGVAPY 720  
ALPGRFLNLY DDGMVIGYS NKLKLVFANK WRGTFWNRH MLVPAFVLF MEVRLKSRM 780  
YSCVNRHSL SQSFQWLIT PQSLICMGL LPLSLIPVG LKQKFPDEL HMYIKELR 840  
IACKRNPFT SCGRFTGLT KLLDSVQPIA RELKQFPFLD LKXSRNVSD PPMHSAITS 900  
VQVPLGKQK VKPIVHRQ

WO 02/098358

PCT/US02/17594

Seq ID NO: 26 DNA sequence  
Nucleic Acid Accession #: CMT cluster

5  
1 11 21 31 41 51  
AGCATATCAT ATGCGCATG ATGATGATG TTGTTCAGGT CTTATCGAGA GTGCTCTGTA 60  
TATCTCATCT CAGCATG ATGATGATG ATTAATCTCT GTATGATGAT 120  
GGAGGCTAAG AGTGTTTAACT TTCTCCGAC TTCCAGTGCT AGTGAATGTT GNNNNNNNN 180  
NNTGAACCTG TGTAAATGCT GTTCTCAGTC GATGCTGTGA TCTGTTGCA CACACTTTGA 240  
ATAATCTTGG ACCTTTGAGG TATGAGAGAC GATTAATAT AACCTCTTGG TATAAATGTT 300  
CTCTCTCTGG CTCTCTCTGA ACATATGAG AAACAGATCT CTACCAATAT TAAATCTGAG 360  
CTATGACAGA GAGTATGAT CATTATGCT TTTTTATG AGAATCTTCT CTGATGATCT 420  
TTTAGTATTA TTCTAAATTA AGCAATTAAC AATGCAACAT TTGTGTCTA AGCAGTTTCT 480  
CTCCAGAAA AAAAAAAA AGTCGAC

Seq ID NO: 27 DNA sequence  
Nucleic Acid Accession #: NM\_06551.2  
Coding sequence: 64..336

20  
1 11 21 31 41 51  
AATTCTAGAA GTCCAAATCA CTCATTGTTT GTGAAAGCTG AGCTCACAGC AAAACRAACC 60  
ACCGAGAAC TGTCGATGTT TTCTCTGCTG GTACGAGCTGG CCTCTGCTG CTACCGAGCC 120  
ATATCCGAGT TGTCGCGAG TCTTCTTCTT GAGCTGTGAG ACTTCTCTCT CATTAGTGAA 180  
CCTCTGTTCA AGTAAAGTCT TGCCAAATTT GATGCGCCCT CGAGAGCTGT TGCGAGCAAG 240  
TTAGAGTGGA AGAGATGCAC GGATCAGATG TCCCTTCAGA AAGCAAGCCCT CATTGCGAAA 300  
GTCCCTGCTA AAATATTTGAA GAATGTGATG TTGTGACATG TAAAGACTTT CATCTGTTCT 360  
TCACTGCTCT TCGAATGAC AATGATGCTT CACTGAGRA TGTAAAGCTT TCAAGCTCTT 420  
GCTTAACTAA ATGCTGTGCT CAC

Seq ID NO: 28 Protein sequence  
Protein Accession #: NP\_065542.1

35  
1 11 21 31 41 51  
NKLAVCLLIV TIALCTGAN ABTCFALVSS LLDFFTFIIEP LFKLSLAKFD APFSAVAALK 60  
GVKSECTQMS LQKSELIAEV LKILKLCSEV

Seq ID NO: 29 DNA sequence  
Nucleic Acid Accession #: NM\_002645.1  
Coding sequence: 1..5061

40  
45  
50  
55  
60  
65  
70  
75  
80  
1 11 21 31 41 51  
ATGCGCTCAGA TATTTAGACA CAGCGGATTT AAGAATGCTT GATTTTCACA TCCGGAACCA 60  
ACAAGACGAA AAGATGTGGA CAAGAAGAA GATTACAGA TGGAGCAGA GCTTTTAGCA 120  
AACTGCGAAA AGGATGACGA AOTGACTGAC AATCAGAGAG GCTTTGAGTT GTCAAGCAGC 180  
ACCGAAGAAA AAGCGACGCT TTATACACAG CAGAGATTAG ATCTCATGCT GTTCTCTGAA 240  
TCTGATTCCT AAAAAGAGAC ATTAGATGAT GATCTAGAAA AGCTCACACA AGCTGAAGTT 300  
GAGAACTAT TCGTGGATGA CATTGTCGAG ACTAAAGAAA CAGCTGTATT ACCAGTACT 360  
CCTATTCTGA GCGCTTCTCT TTGAGCAGAG CTCATTATT GACTACTAT TCAGAGAGGA 420  
CAGTGGCCAC CTGATATACC TGGGCTCTCC ACTATGCTT TACTCTGAT TATGCTCTGT 480  
ACTTACGATG AAGAGCGCGT ACTCCAAAT GCTATGATCT GAGAGATCC CACTTTCTG 540  
TCTCAGAAC CATTAGATCT AAGTCTCCG GGCACATCTC CAXATCTCTC ATATCTCTCT 600  
ACACTGCCA CAGCCTTTCA TCCACAGGA AGCTTACCTA TCTATGCTCC AGTAGAGT 660  
ACTGACATGG CAAGATCAT TTACAAATA GCTAGTACAT CAGATATTT AAAAATGGG 720  
AAGACAGGA CTGATTTGGA GATACAGAT TCAAAAGTCA GCAATCTGCA GGTATCTCC 780  
AAGTCTGAGG ATATCACTAA ACTTGAATCG TTAGACTTGG ATGCTCTAG TAAAGCTAAG 840  
GTGATTAATG TGGAGTATT AGACATGAG GAAGAGAAA ATGTTCTGAG TTGTGACGA 900  
AAGATCCTT GGAATGCTGT TTCTCTGAAA GAGAGATGCA CAGCAATGTT TATCTGTGAA 960  
AGGAGATGGA ATGGAATAAC CTCTCTCTGT CACAGATCTA CAGAGAGCCA GCTCTTAAT 1020  
ATGATGAGA TTAGCTCTCT AAAAGCTCAG GCGCTGAT CTACAGAGA CTCAATGAG 1080  
ACCATGATT TGCCACATGG AAGTCTCTCT CTTCAAGAG TTGAGTACA GAATGAGAG 1140  
ATGCGACTGT TTGTGATGCT CATTACAAA AATTCCGATA TACCAATCAC 1200  
CTCACAAACC CAGCTATTCT GTTAGTCTCA TGCACAGCC AAGAAACAT ATGCGGAGAA 1260  
ATGCTGATG TGAAGTGTCT TTATGACATP GAGGATATTC AGTACACGCT TACTTTTGG 1320  
TGTGATGTA GTTCTGATGT AGAAATCAT ATATGAGAG CCGTTCCTG GGTACATGAT 1380  
GACTTGAACT AATTAGATTT TGGCAGTAT TTCTTAAAG TTGTGCTCA AGAGAGAGT 1440  
CTGCGATGTA ATCATGCTCT TGGAGTGTG GAGCATTTAG AAAACTGTGG AAATCTGAC 1500  
ACGGAATTA GACTGATCT CTGATCTCTT AGTCTGATCT GTTAAATCT GGTCTGACA 1560  
CGAGAGATG ATGACAACAC CTGGAATT TAAGCAACAC TGTATCAAT AGAATACT 1620  
TCCAAAGAG CCGATACGAG ACACCTTGT GAGATCTCT TAGATTCTTA TCACAACCA 1680  
GTAGATCTGG CTCTTCAAT TGAAGACCA CAGCTGCGAG TAGATCAAGT AATTAAAGCT 1740  
GTAGAGATA TCTGATGCTT TTATGATGCT TTGTGATCT TTGCTATCT AGATCATCT 1800  
AAGAGATGTA AAGAGACGAG TAATCTTCCA AGAGATTAATA CTCTGATGTT GACTCTTGG 1860  
TTTGAGGAG AGACACTAG CAGAGATGCA ACTAGGGCTCT CACTTAATCC TGAAATCTCT 1920  
CTTCAAGTAA GCAAGACCA ATGATGCTCA GGAATTATG ATCTCTGCG ACTCATCTCA 1980  
AATTTCTGTA GAGATCTCT AGATGTTGTC CAGATCTGTA AGATCTGTA GCGATCTGG 2040  
ACTACACAG AGCGCTCCA CTCTACTAT TTGTGCTGCT ATGGAATTC AGTAAATGG 2100  
GTATCAAAAT ATGAAATAA CTACTTGATA TTGCTACTGT CTCAAAAG AGAAGATCT 2160  
TTGAAOCTA TTCAATCRAA GAAGTTTGG ACTTACAGA ATTTCTTCTA TCTTATTAAA 2220  
TGGATGAC TATCATTTT TCTATCCAG ATATCCGAT TGCATGGA ATCGTCTCT 2280

WO 02/098358

PCT/US02/17594

5  
10  
15  
20  
25  
30  
35  
40  
45  
50

CACCTTACTC TTTTGGAAAT TTAAATACAG AGAGCTGGGA GTTCCCCGGA TCTCAATAG 2340  
CAGAGAAGG GACCAAGAAG TTTGGGAAA GTTCTTTCAT CTCTTGTGA CTTPAGAGG 2400  
TTTTTAACAT GTGAGACTAA ACTCTATAT CTCTGACTC CATCTCATAC AAATTCCTTT 2460  
CTCGACAGT TTACCAAAA AGCATATATC ATGGAAGAAA TATCTGACGA GGTATGATTT 2520  
CTCTCTCCG CATTGATAT TTTTATTCAT ACTCTTCAG TTAGACAGAG CATATACAG 2580  
CAACATACT TAGAAAACAT AGGAATATAT ATAAANGGA AACCTCTTGA TATCTTCAT 2640  
AAGAATCAT CACTTGACT TCTTAAAGAA ATAAANGCT TTTTATGGG GAACCTTAT 2700  
TATCTGCTCA AAGACCCAAA TCTCTTCTCT AAATATATAG CAAAGCCGCC AAATCTGAAA 2760  
TGGGTATCT TTGCAAAAC TATCTGATCT CTTCAGATC AGCTCTGAT TCTCCACGA 2820  
ATTGCACTGG AACCTCTTGA TTTCAAAATTT GCATGATCAG AAGTAAGATC CCGAGCTGG 2880  
ACCTTGATTT AGGCACTTAG TATGATGAGC CTACAGATCT TCTCTGCACT GTTTGTGACA 2940  
CTTTTGAAAT ATGAATTTA CTGGAAGAGT TCAATAGTGC AAATCTCTTT CTTCAGGCA 3000  
TTGGGAATA TCTGCAAGC AGCATATAT TATCTGATCT CTTCAGGCA CCGTATGAT 3060  
GTAGACTTIA GTACCCGATA CGAACATGTT TTGGGTGCTC TCTCTGAGT AGAGGAAAA 3120  
CGACTTAGAG AAGACCTTCT AAAACAGAGC AAACCTGTAT AGCTTTTAGG AGAGTAGACA 3180  
GUAAGATGAA GGGGCGTAG TGGATGCCCT AGACAGGCTT TCTCTGAAAG AAGTAGGAA 3240  
CGAGTACAT CCTTTTTCGA GAAAATGAA TGGCTGCTC CTTCAGGCA AAGTATGAT 3300  
GCAAAAGAT TAAATATGAA GTGCTGTCTC TCTCTGAGT CTATGCTGCT CCCCCTAAA 3360  
GTGCAATGTC TGAATCTGGA CCGCTGCGGA GAAGAATTA ATGCTATGTT TAAGGTTGGT 3420  
GAAGTCTTC GCAAGATAT GTTACTGTTA CGAATGATA AAGATATGGA TAAGATCTGG 3480  
CTTAAGAAAG GAGTATGCT GAGATAGATA ATTTGCAAT ATCTCTGAC TGGCAGAT 3540  
CGAGCTAGG TGAAGCTGGT TCTGCTCTC GATCACTCCA GGAATACTCA AOTGGAATAT 3600  
GGTGTGACAG GATCTTTTAA AGATAAACA CTTCGAGATG GGTCAAGGA ATACAATCTC 3660  
TCTGAAGAAG AATATGAAA GGCCTGAGAG AACCTTATCT ATCTCTGCTC TGAATGCTGT 3720  
GTAGCACTCT ATTTTGTGAG CATCTCTGAT CAGACAGATC ACGATATAT CTCTCGAGC 3780  
ACCGAGCACA TGTTCACAT TGACTTTGGA AGTTTGTGG GACATGCGA GATGTTTGGC 3840  
AGCTTCAAAA GGAATCGGCG TCTTTTGTG CTGACCTCTG ATAGGCGATA TGTCAATAT 3900  
GGGGGTGAAA AGCCACCAT TCTTTTTCAG TCTTTTGTG ACCTCTGCTG TGGGCGCTAC 3960  
AACTGTGAAA GAAGACAGAG AACCTTTTCT CTATACCTCT GACTCTCTGAT GATCTCTGAT 4020  
GGGTAGAG TATTACAGT TATTACAGT TTAGACATCT TTAGACATCT ACTTCACT 4080  
CAAACTACAG AGCTCAAGAG TACAATTTTCT TTACTAGGCT TTATGTATCT AAGTTTGGGA 4140  
AGCATTGCCA CAAGTTTAA CTCTCTCAT CTGACCTCTG CTGACCTCTG CTCTCTGCT 4200  
GATCTCTCA AGAGAGAGG GATCTCTCTA TTTTACCTCA AAGATATCTC CTTTAGACA 4260  
CTTCTCTCA TCAAGACATCT GATCTCTCTT ACATCACTC CTCTCTGCA 4320  
CAATCACTTT ATGATGCTCG AATTTTGTG GAAGACAGGA TTGACCATCT ATTGTCTTC 4380  
CGAACATTGG TGAATTTTCA GGAATCTTCA AATAAGCTCA GTATTATTTT TCACTTTGG 4440  
AAGTATCAGG GCTTTCTTAA TAGAGTGCTT CTAGAGAGGA CACATATATA AAGTATGAGA 4500  
GGCAAGGA AATCTGCTCT AATCTGCTCT CTACAGATCT CTACAGATCT 4560  
GTAGACAGAT ATGATGCTGT TGTGACTTTC TACTGCTCT TACTGTGGA TGAAGAAGCT 4620  
GAAGGATAG CTAGCTCTCG AGATGCAAGT TCTCTGCTCT CTACTCAGG CCAATAGGA 4680  
GAGCTGTGGA AATATCTCAT CTCTTACGAA AATGCTGCTC TTTTCACTAT GAGATATGAT 4740  
ATCAAGACT TGTATACGA AGATGCAAGT GACCTATATC CAGATGTGA AACTATGCTA 4800  
CTTCCAGATA ACCACAAAC ATCAAAAGT AAACCAAAA TTTCCAGAAA AACGAGGAT 4860  
CGACATTTCA ATGAATGCT TGTATACAGT GATATAGCA AAGAAACCTC AAGACAGCA 4920  
GAATCTGAC TAGGTATGCT CAGTGCGAGA TCTCTGCGG AAGATTTTTT CTGGGTGGA 4980  
GTAACTCTCG CTGTAAGGA TTTCACTGTC ACGAAGAGA CGTTTAAATG GTATCTGAGT 5040  
ACTCGGCA CAATCTTGA A

Seq ID NO: 36 Protein sequence  
Protein Accession #: NP\_002636.1

55  
60  
65  
70  
75  
80

1 11 21 31 41 51  
NQCFTSNGSF KCPFFHPPEP TRAKDVKEE ALQHIAEALA KLGKDRQVTD VQGPFELSS 60  
TKKQAQYVK QDTDLAVPE SSGSKALDI DVEKLQAEK KELLLODSFR TKCTPVLPT 120  
PILSPSPSAQ LYRFPTIQG QWPFGLPGS TYALSFYPS TYSQEAARON GFNPFPTFP 180  
STEPILYSLP QSDPYFSLP TPATPHPGQ SLPYRFPVS DMAKLDPKI ASTSEPLANG 240  
KATCLGTTD SEVSNQVSP KSEHISKFW LLDLDELSP VNAVSLHSE SEKVSLSLA 300  
EDPKVLLS SEATACVLE KVGSELGV AYTRSGUL IETDLAAG HELSDRNG 360  
TSLPLTGSLL LEQVEVQHE MVAFCSITK LTKTFPYVNH KTNFGLLSE VTAGRMIOGE 420  
NASVKVSDI ESQLPVTTF CVHSTVEI I DAQALGVHD DIANQVGSY VLKVCQREVE 480  
LQNMICGSH BILGNCKRWG TEIRLGLLPT SAKCKMLAT AGDDSTVDL SHHLLISRP 540  
CKSANTHPY HILGSLQVQ HELALQIENQ HADQGVLIA VTKICCALG VETLALTSY 600  
KELKAVNPL RSTADVSLP PGEIDTSSS TROSLNPNP VQVSEQLTA AYLDLKLIA 660  
NEGRSPDCA QSSESVKDA TTEQLQPTI PAHGISNHW VENTKYVLI CSLSINGKDL 720  
FKDQSKNGY TYNFPLIK WDLILPPIQ LQPLFSLVLT HLLPLGLMG SSSSPSPSK 780  
QKSEHLEAK VSLGLQVSP FVSGEDALV LTVSESTINW PPTVTKSEY HSRVLVQD 840  
PEPAFDITIT TQVDSRSLQ QMLTLELND HSKLLDLIL KDSLSLGEK DKAFLAKGY 900  
YCFHPNCLP KLSAPHWK WNLAKTYSL LQNPALPYL IALRLDSKF ADQREVALAD 960  
THIDAIDE DTLALPQVQ ALRYEITLMS SLVQFLERA LGNIQIANG VHLDELIED 1020  
VQFTIRVBL LQMLALVQV ELRLGLLQGT SLVQLAGRA EYVAGASA RQVVLKRG 1080  
RVQSPFDKQ CLPLKPSLW AKELIKSCS PFSNAVPLK VTMWADPLG SEINVPKVG 1140  
EDLRQMLAL QMIRIMDKIN LKESLDLRNV IFKLCLTGRD RGVSLVPSA DTLKIQVEY 1200  
GVYSPKQKP LAMKARKVP SESEYKASB NFVYSCACC VATTYGLICI RHNDRHLS 1260  
TBNHIVLQV KFLGLQVSP SPYSDRAVPL LYSNVAIVN GSEKFTIRV LTVLQCDY 1320  
NLISQNTNF LNLGLSLPFS CLPLTSIQ LKTVEDALQ CTTDNATFI PQLLSSLSG 1380  
SIATKPNFTI HNLGLRFSG LPSNDEPLS PSKTFSPRQ DGRKEVSFV THKKYKDV 1440  
HITVIVKIL BOQIESFVFP KTFVBEQML NGLSITPLW KLPGFPRMW LKTHKEDVA 1500  
AKSKIKARI LQMLNATO VASDDEPLT FHPKLCVST SIATKPNFTI SPYTPDQIG 1560  
GAVKLSYR NQPLFMVNI IDLVYEDGA DMVYVKVYL LFNQNSKRS KTKISKRTN 1620  
PTPHMLVYS VSGKTLQAR ELQLSVLGS SLBNPFLAG VTLPLDPLH KEIVVAKQL 1680  
TAATVL







5 TMAAATAGA AAGAAGAGA ATAAAGCTTA GCTCTGTGTC TTAAAAATT AAAAATTTTA 4500  
 CTTGATCCOC ATCTATGCGC TTGAGACCTA TTACGTGCGT GAGTCTTAA GTTATAATTG 4560  
 TTCAATATGT TTTTGAACA GTGTCGTAAA CCAATAGCAA ACCCATGCC ATATTAGTTA 4620  
 TTCTCAATAT TCAAAAAA TCCAGCTAGA TTGCGATTGA ATGATTAAAC TGCACAGACT 4680  
 GTCGATATGA TGAATCTTGA TCTATATGTA ATTAATTTTA GAACCAAGAT TGGGAJAAT 4740  
 GCCTCTGTTT CATTCGTTT AATTAAAGCT ACCTCTTAAA CTATAGTGGC TGCAGATAGC 4800  
 AGACTGTGTA ATTCGGTITT ATATACCTTT TGCATTGTAA ATAGCTTTTG TTGTAACCTG 4860  
 TCGGTGTAAI AAAAAAGGA TCTTTGTATTA TCAAAATCAT TTAGTTCTGA TAAATATGCG 4920  
 GAAGATATTA TGAATCTTGA TGTAGAGCTT TGTAGAGGCT ACTACTTACA AGTTTATAAA 4980  
 10 ATTCTATCA TATATATTA CACATCTGAT AAATATTAAA TGAAGACTGT GTAGAGAACT 5040  
 CCTATTATTA AGGTTTTCCT CAAATCTCAG GTATTGTAAA ATTTTTCATT TTATTCTATT 5100  
 AAAAACTAGA ATACAGATA TATAAAGGTG TTATCTCTTG TGCATATG TGTAAGAAATC 5160  
 AATATTGTAC TCAAGTGTTT GAATTATTGA AGTTCTTAGA AAGCAAAAA A

Seq ID NO: 36 Protein sequence  
 Protein Accession #: NP\_060960.1

20 1 11 21 31 41 51  
 NPGFLGLLCT LALGLLGGAG P9GAAPFLCA APCSCDDRR VDCSGKLLTA VPBGLSAFTQ 60  
 ALGISRNITL QLEDAFDFP PFLBGLZLAG NULSPHMPKA LSLRLKLV TLQRNQLRTV 120  
 PSIALRGLA LQSLRLDNR ITTPRDETFP GLVQLRLHLL LNSLTFVFP HPLSLPLTLQ 180  
 ALTALNKIS SIYDFATNLL SSVVLHLLHN NRIKLGQHC FGLDNLBTL DLSYHNGEP 240  
 PQAIKAPSL KELPFRNSZ SVIYDGFDP NPLRLTHLY INPLSPVNGS ASHNLSDIHS 300  
 LVKRGASVQ QPRLATVTH LSLTLATNKH ISSTPMLCC KQMLKTLGL SPNNHKLPS 360  
 25 FVDCIALSLI EAPQVQVGT KPTPTQLS LRLGLBWL CHHEHBAFA TLPITNLV 420  
 SPNLSLFTPT EAPGLQLQL LVGNFKLESA LAACFPNLR SLSPVYAIQC CNPNCDCYKA 480  
 MLSTEDSLQ DHSVAGEKT ADARNVISTL EHSEHGOIII KCTSPSTGAFK PCYLLGSMW 540  
 30 IRLTWFIPLF VALFPHLVZ LTTFPACTGL P6KRLIGLI SVBLNPNLYI TGLTFPLDAP 600  
 SWKTFARLTA TGAATLVAFVSE LKTLFLLAL PYSKLAALH KSGRNSMLH 660  
 QFWAALSAF LSAIVACDFP LFRGEYBBS PLCLFPFPGT TPLGFTVTL VLNSLHAP 720  
 NAVITKLYC NLEKEDLSN QS8SMIKFA WLIFTPCIEF CPVAFSPFAP LITATISPS 780  
 IMKSVLTFP PLPACAPVL YVFPNPFKE DWLKKERVY KSGSVSVSI SQQOCLRQ 840  
 35 FYDQSWISL LQSLTVDCS DQVGLTFEFP SKHLISHS PLALVAQAG RFDGWSGDC 900  
 TQMSHSDAC EHSFVSUS DQVQACBAC FQSGRFFLV KTAHLPRVK D

Seq ID NO: 37 DNA sequence  
 Nucleic Acid Accession #: AF144648.1  
 Coding sequence: 1..1884

40 1 11 21 31 41 51  
 ATGCTCGGAG CCGAGTGTAT CCGCGTGCCT ATCAGGACCT GGTCTGCGGA GGGCAACTAC 60  
 CCGAGTCCCA TCGCGAAATT CCACTGTGAG TTCTCCTCTG CTGTGCGCGA AGTGTCTCTG 120  
 45 AACCTCTTCA ACTCGAAAAA TTGTGCAAT GAAGCTGTGG TCCAAAGAC TTGTGACAGG 180  
 GATGCTGCA GAGACAGTPT CCGCTGTAGA CCGAATTGTS GAGGTGCGCT GTGTCTCTG 240  
 AGAGATATA TTATGTGAC GAGACATGAA CAGATCTCAG AATGAAATAT GAGATACAG 300  
 ATCAAGTATG TTTTTCATCA GACTTGAGAA GATTCAAGCT TAGCATCTA TGAAGCACCC 360  
 CTGACCTTGA CCGTGACTA TCGGATGCAT GAGAAAGTGT GGTCTCGCTGA CTGTACTTT 420  
 50 TGAACACGA AGATGTCTT GTGTGATGAT GTGACTGTGG AAGATCGGT GTTTCAGCTT 480  
 CACGAGATAT GAAGGTGGS GTACCGCTCT CAGTACACA CTACGACAC TTGTCTCTG 540  
 GATCTGKATA AATCCCTAT GACAACAGAG GCGTCAACCC TGTGTGTAGA GAGCTATGG 600  
 TACAGGTTG AAGACATCAT ATTAATCTGG GATGACAAAT GGAACGCCAT CCACTAGACT 660  
 55 CAGAGACTGC ATATCTCTCA GTTCACTTTC CTGAGAAAGA GAATTAAGAG CAGAGAGTGT 720  
 TATTCTTCA CAGTCTCTCA CACTACGCTG ATCAGAGAT TCAAGATCTA GAGGAAAT 780  
 AACAGCTACC TTGTGACAGT CTACTGCGCT ACTGTCTCCA CCACTATAC CTCTGTGATA 840  
 TCGTTTGGG TGAACATAGA TTCTCTGCA SCAGAGGTGA CATTGTGCTT AACTTCAATG 900  
 CTCACTCTGA CCAACATGTA CTACATCTGG CCGATTAAGC TCCCAACAT TTCTGTATC 960  
 60 AAGCAAGTAT ATATCTATAT CCGCTGTGCG TTCTCTCTHS TGTCTCTGCT CTCTGTGAG 1020  
 TATGTCTACA TCACTATCT TTCTTACAT CAGAGACCTC GGGCGAGCC TGGCGACAC 1080  
 AGGAGACCCC GAGAGTCACT TCCGCCCTAC CGTCAACGAC AGTGTGTGT AGGAAACGTG 1140  
 CAGAGTGCCC TGATTAAGT GSAAGAGGGA TCGAGCTCTC TCCCACTCAT CCGACAGCG 1200  
 65 GCGCCCGTGG CAGAGCCGGA AGACAGCTGT TCTGTAGCT CACCTCCGGA CAGAGCCGAG 1260  
 CTGCGACCT CCGAAGCCCT CAGCCCATCT GACTCTCTC CAGCCGAGC CCGCTGCC 1320  
 ACTGAGAGAA GCGTAGACGA TCTCCCTCTC ACCTCAGAGC AGGCGCGCA CAGCTATGGT 1380  
 GTTCTCTTGA ATGTTTCCA GCGTCAACAG AGTATTTTTC CTACCGAAT CCGCAACCT 1440  
 70 GTCAGAGCC ATGGCATATG TGTATCCAT GACATAGAG ATTCATAGA GAGCTGTAGC 1500  
 CTGAGAGAC GCGTAGGCA TGGCCCGAG GGAAGACCA TCTTCTACA TGCACAGAG 1560  
 GTGTTCAGAG AAGCAGCTGT GAGACTTGAT GACACAATG ACAGAGACGA CTGCTTCC 1620  
 ATGAGGAGAC AATTCAGAG TGTACTACAC AGTACTCTGG CCGTTATGA TGAATGACTC 1680  
 ATGACCCATG GCGAGAGGA AGACATAGC TCGAGCTCTG AGGTAATGT CCGCCACAG 1740  
 75 CTTGCTCTCT CCGTACATA AGGFTCTCTC TCGATCTCT TCACTCTCT CACTCTCTC 1800  
 AAGGTGACA AGTGTCTCG GTTCTCTCTC CTTCTGCGCT TTGGTGTGTT CACACTTGT 1860  
 TACTGTGAT ACATATGTA TTAG

Seq ID NO: 38 Protein sequence  
 Protein Accession #: AAD51172.1

80 1 11 21 31 41 51  
 MRAAVILLI IETWLAGNY FSPIPKHFPE FSSVAPPEVL MFPNCINCAN EAVVQKILDR 60  
 VLSRYDVLRL PTFGQAPVY RLSIYVTSIB QISENNDET ITNFIFITKH DRLATYET 120  
 LNLITLDRIH EKLWVPCYF LMSKAPVMD VTVERVQPL RHDSTVTKG RLTTTAACSL 180



TCCAGGTCCC ATACGTGAGA TATGATGAGG ACTACCGAGA ACTGATTGAA GACATAGTTC 2880  
 GAGATGGGAG GCTCTATGCA TCTGAGAACCC ATCGAGAAAT ACTTAAGGAT AGAGAACTTA 2940  
 TCAAGGCTCT GTTTGATGTC CTGGGCCATC CCGAGAACTA TTTCAAGTAC ACAGCCCAAG 3000  
 AGTCCCGAGA GAGGTTCGA AGATGTGCTA TCGATGTGCT ACTGTCCGAA GTGTCCAGGT 3060  
 TTTTGAGAGC TATGCAATGA CTGAGCCGAC GTCCACACTA ATGCAAAATG TCTGCTATAG 3120  
 GATTGGTGGG ACAGAGCTGT CTCTCTCTGT CATGTGAGCA GAGTGCGGTA TTGCTGCCCTC 3180  
 CCATATCGAT GACTGATGAG AGTTGAAATT TTATAGATAA TACAGATATT TGGTAAATT 3240  
 GAACCTGGTT TTCTTTCCTC AGCACTGTGG AGTATAGAGC AGATAGTGCT TGGATGTCAT 3300  
 GAGCTCTCA CTGCTGGGAG CGATGCTGCT GATATAGACT CAGCTGAGCT GACTCTGACT 3360  
 TGGTGTGGCC TGTGTGAGAC TGCATCTGCC TTCTGGGGTC TTACTTCTCT TCAAGAGTCT 3420  
 TGTATGGGAA AGGAGGCTAC AGAATAAGCT GCTTATTCTC AACTCTGAGC TCTCTCTAGC 3480  
 CCGGCCCTCT TCAAGGAGAG CTCTGCGACT GTGTGTGAGG CCTTGACGAG GCGAAGAGGT 3540  
 CAGAGGAGGA GGAAGAGGAA CCTCTGAGG CTCTCTGAGC CACTGTGAGC ACTGTGAGG 3600  
 ACTGAGTTTC TCGGAGCTCT TCTCGAGCTC GTGTGATAGA AGTTTGAATC CAGGAACTTG 3660  
 AGTTCTAAGC AGTGTCTCTG AAAAAAAAAA GCAGAAAGGA TTGAAGATAA ATAAAACTTA 3720  
 AGCACTCTGT GAGACAT

Seq ID NO: 42 Protein sequence  
Protein Accession #: NP\_066025

1 11 21 31 41 51  
 | | | | |  
 MTVAGNRPGR AAKAVLLLLL LEPFLLLLAG AVFPGRGRAA GPQBEVDCEA QQLDDCHADA 60  
 LQNTPTSVK CCKPKYQYQG GRQCEIDIBC GNELAGGCVH DCLNIPGNVR CTFDGPMLA 120  
 HGGHCLDVD ECLINNGSQ HTCVNWSBY SCCKKBPFL SRNGTCHIR RRLGLGQNR 180  
 DISCHLICH APGIRAGCG RPTLEAARG RDCILPHWG KQCHGECQ TADPDCSC 240  
 PLYKMDIGDR SCLEBEDTVL EVSTENSTV VGGKVRVER LAMETCALNV GCDRTCKDT 300  
 STGVHCSCPV GPIQLDGKRT CEDIDECQTR NGGCDHFCGI IVSGFDGCKK RSKFLKLTDEK 360  
 SCQVDEBCEL DEPCDISCHN HPGTACACAN KGTTLTGPHI CQPNHCSIE NGGQYQVGN 420  
 TGTGAGTCA PWTGAKRLLF VSPVSRVGLC TSPVSRVGLC PLRCHSCG 480  
 BSVDTTICS VTFELNCKEK SLNKAELFPE GLAPLAPENI SVKSKPRV MLTCSBKQV 540  
 PGAPRSPST KEMFIVTFEE LETNKEVTA SCILSCIVKE TSKLEKKAIR TLKKAVERBQ 600  
 FILOLQSGOL DVAKPPFRT BRQASGVCG QGMAHQVCB CRAFTYIDGA REBCLICBQG 660  
 FQREDSQMT CEFKAPRBS HECGLGGLG HECGLGGLG CQGLALGPT CQGLALGPT 720  
 FENRTSPF CQGLALGAGG GATSPQDCBT RVQCSQWVE NTHTHRCIR FWTYQPEFS 780  
 IJNVCSCVPN TTDPDGNTN ITQCKNRRCO GLBIDPTVI BSNVPHNPT ANTECTWIN 840  
 PPKRRLIVL VEFIFLEID DCDYLVNMX TSSSHVTVI ETQYTVBRI APTSRKELW 900  
 IOKSNERSR AGQFQPVVT YDFYQELER DIVYDRLHA SRHQSELED RKLKALFDV 960  
 LARPQVTVY DQSSSEMPF RSPFILLKX VSEFLPYX

Seq ID NO: 43 DNA sequence  
Nucleic Acid Accession #: CAT cluster

1 11 21 31 41 51  
 | | | | |  
 TTTCTGAT ATGAGCTTTT CTCCCTTTTA TTTACTGCG GCGGTTTTC CTGCGAAAT 60  
 TTTCCCTCTC ATTAATCTTC CTGACGAAT TCACTACTCT TTGGGGGNTT CTTTGTGTT 120  
 TTGATATGA CTACTACCTC ACTGTATATA GTTCCCTTCT TTTTTTTTC CTCCAGATT 180  
 CTCTCTCTTC TACTGGKATC CTTTTCATT TTTACTCAAT TTCTCAOTT AGTTGACTT 240  
 CTCTTTATAC CTGTGTGATG TCTCTTCCCA GATATATGAC AAATGCCGCC AGGATCCAT 300  
 CATTTTCTTC CTAGAGAACG TCGAAGAGAG CTTGTGCAAT AACAGACTT GGAATAATGG 360  
 AATCTTTAA TATACAAAGC CAGTGAATTC TACTTGAGAG CCAATGCTTA GAGGCAAGG 420  
 ACAGTGATCT AATAGGTGT TGANNNNNNN NNNNNNNNN NATGATCAGC ATAGCAAGA 480  
 CTCTTTCTCA AAGATGAGAA GTTATGCTA TTCAATATTA OCTAGACCTT TTAACAGCC 540  
 TTAAATGAT TAAAGAGCA AGAATTTTAA GTATTGAGA ACTGAGTAT AATTAAGAG 600  
 AATTAAGCT GTTTTGAGG GAGCACTGAC CACTGTGATT GACATGCGG GCACTTACTG 660  
 TTAAGGGAT TTATACAGA AGTACTTGAA CAGAAATGTT AAGAAGATAG AATTGTGAT 720  
 TCTTTTATCT GCCCAAGAAC AGAGCTCCCA TGGGAATATC TCAACTCAT TCTACTACT 780  
 TCTGCTGCA ACAGACGAG TCAATTTTAA ACATATCTCA AAGGDTTATC CTACCCGAC 840  
 TTGAGAAAT CATAGCATNC TCCCTTTGAC TATAACTTGT TCCATATGAA ATACATCAA 900  
 ATGCTCT

Seq ID NO: 44 DNA sequence  
Nucleic Acid Accession #: CAT cluster

1 11 21 31 41 51  
 | | | | |  
 TTTTTTTTT TTTTITGGA TTTTATGATG OCTTGCAATT TTTTCCCTT ATCTGTAGTC 60  
 ATGAGTACC CACTAAAAAT GACTGCTGT AGTATAGCTT CAGTAATGAG GTGATGAGT 120  
 GAGGAGGAGG GTATGCTCT CTGAGTCTCT TTAGGCTACT ATTAAGAAAT ACTTOMAGCT 180  
 GATCAATCA TAAACAGG AGATATGTT TCAAGGCTT CAGGCTGAGA AATTAAGAG 240  
 CTAAGGACC AGAATATTG GTTTTGTGTT AAGTCTCAAC ATTCAAGAC TCTCAAGAC 300  
 TATAGGACA CAGACGCTCT TCGAGAACTC TATGTGAGG ACAAAGCTC AGAAGGCGC 360  
 TCGAGTGTTC TAGATCTTA TTGAGGAGAC AATCACTGAG AACCAGACAG TAGTGTGTG 420  
 GATCTCTAT TGAATGAGA ACTTGTGAGC OCTGTGAGA GTGTCTGAGA ATCTATGTC 480  
 AGGACAAA ATTCAGAAC TTGTAGCAT GTTCTGAACT OCTATGTGAG CACAATCA

Seq ID NO: 45 DNA sequence  
Nucleic Acid Accession #: Bos sequence  
Coding sequence: 31..1092

1 11 21 31 41 51  
 | | | | |



WO 02/098358

PCT/US02/17594

	TCCTTTTCAG	TCATCTACCA	CACACACAGA	CACACACAGA	CACACACAGA	CACATATATT	240
	TATATAGTGG	ATGGGACAGT	GTGCTACGTT	ATATATAAAG	GAAGTAAGGT	TTCTTTTGAG	241
	ATGAAATGAT	TCATAAATAT	TCATCGGGCA	TGCACATCTC	TGAAATATTT	AAAAGCCATT	360
5	GAATGAGAAA	AGGCTGGGG	AGATGACGAA	ATGATGAGCG	ACCAACCTCT	GAAGATTTCG	420
	TTCTTGAGAA	TGAAATTTCT	ATATACATGA	ATGATGTAAA	AGATCACACC	TTCTCGAGGT	480
	AGATGACGCA	AGATGACGCA	AGATGACGCA	AGATGACGCA	AGATGACGCA	AGATGACGCA	540
	TAGACATCTG	ACATCTATCA	ATTTTGGGCT	GTCTACTACT	AGTGTCTACA	TAGGTATAGT	600
	GAAGAAAGTG	TACCTTGATCA	ACTCTGTCTG	TCATTACTCA	GAAGAAATTT	GCCTCTTAAG	660
10	GTGATATCTA	TTCTTGCTGA	ATGATTCAT	GAATTTTATT	TACTTCATTA	TTCTCACAGC	720
	AGATGACGCA	AGATGACGCA	AGATGACGCA	AGATGACGCA	AGATGACGCA	AGATGACGCA	780
	ACAGAGCTCG	TCATCTGTATC	AGAGAAATCA	CATGAAATAT	GRACCTGTAC	TTTTTGTATG	840
	CATATGCTAT	TGACATCTAT	TGAGATATCT	TTCCTGCGA	TCTTACACA	CGATTAAGAG	900
15	CTCATGTAGA	AGATGACGCA	AGATGACGCA	TTAAGATGAG	GAAGAAATCA	ACTGTCCGAC	960
	AGATGACGCA	AGATGACGCA	AGATGACGCA	AGATGACGCA	AGATGACGCA	AGATGACGCA	1020
	GAAGCAAGAG	AGATGACGCA	AGATGACGCA	AGATGACGCA	AGATGACGCA	AGATGACGCA	1080
	TGGATCTGAA	ACAGCTGTCT	CTAGATGCTT	GGGATATCTG	GAAGCTGTCT	TGACCAACTT	1140
	TTTCAACAAA	CATCAATAGA	AGATGTATG	AGGACATCTC	GAAGCTGTCT	CGAGCTGTCT	1200
20	TTCTTTCTTG	TTCTTGCGAG	AGATGACGCA	AGATGACGCA	AGATGACGCA	AGATGACGCA	1260
	TTGACACAGA	TGCTCCACCG	CAGGCGGCGC	AGATGACGCA	AGATGACGCA	AGATGACGCA	1320
	CAGCTCGAAG	CTCAAACTG	GGATTTGCTC	ATTAATAGAG	AGCTGATGAG	AGAGAGGATCA	1380
	TTCTTGCGCA	CTAGCTGCTG	CTAGCTGCTG	CTAGCTGCTG	CTAGCTGCTG	CTAGCTGCTG	1440
	CTAGCTGCTG	CTAGCTGCTG	CTAGCTGCTG	CTAGCTGCTG	CTAGCTGCTG	CTAGCTGCTG	1500
25	AGAGAAATGA	TTCTTAAATAT	AGATGACGCA	AGATGACGCA	AGATGACGCA	AGATGACGCA	1560
	CACGCTTTTG	ACTGTGCTG	AGGAAATGAT	AATGTTGAAA	ACTATTAATAT	CGGCGGAGC	1620
	TTCTTATCTC	GGATTTGCTC	CGATGACGCA	AGATGACGCA	AGATGACGCA	AGATGACGCA	1680
	AGATGACGCA	AGATGACGCA	AGATGACGCA	AGATGACGCA	AGATGACGCA	AGATGACGCA	1740
	AGATGACGCA	AGATGACGCA	AGATGACGCA	AGATGACGCA	AGATGACGCA	AGATGACGCA	1800
30	AGATGACGCA	AGATGACGCA	AGATGACGCA	AGATGACGCA	AGATGACGCA	AGATGACGCA	1860
	AGATGACGCA	AGATGACGCA	AGATGACGCA	AGATGACGCA	AGATGACGCA	AGATGACGCA	1920
	AGATGACGCA	AGATGACGCA	AGATGACGCA	AGATGACGCA	AGATGACGCA	AGATGACGCA	1980
	AGATGACGCA	AGATGACGCA	AGATGACGCA	AGATGACGCA	AGATGACGCA	AGATGACGCA	2040
35	AGATGACGCA	AGATGACGCA	AGATGACGCA	AGATGACGCA	AGATGACGCA	AGATGACGCA	2100
	AGATGACGCA	AGATGACGCA	AGATGACGCA	AGATGACGCA	AGATGACGCA	AGATGACGCA	2160
	AGATGACGCA	AGATGACGCA	AGATGACGCA	AGATGACGCA	AGATGACGCA	AGATGACGCA	2220
	AGATGACGCA	AGATGACGCA	AGATGACGCA	AGATGACGCA	AGATGACGCA	AGATGACGCA	2280
	AGATGACGCA	AGATGACGCA	AGATGACGCA	AGATGACGCA	AGATGACGCA	AGATGACGCA	2340
40	AGATGACGCA	AGATGACGCA	AGATGACGCA	AGATGACGCA	AGATGACGCA	AGATGACGCA	2400
	AGATGACGCA	AGATGACGCA	AGATGACGCA	AGATGACGCA	AGATGACGCA	AGATGACGCA	2460
	AGATGACGCA	AGATGACGCA	AGATGACGCA	AGATGACGCA	AGATGACGCA	AGATGACGCA	2520
	AGATGACGCA	AGATGACGCA	AGATGACGCA	AGATGACGCA	AGATGACGCA	AGATGACGCA	2580
	AGATGACGCA	AGATGACGCA	AGATGACGCA	AGATGACGCA	AGATGACGCA	AGATGACGCA	2640
45	AGATGACGCA	AGATGACGCA	AGATGACGCA	AGATGACGCA	AGATGACGCA	AGATGACGCA	2700
	AGATGACGCA	AGATGACGCA	AGATGACGCA	AGATGACGCA	AGATGACGCA	AGATGACGCA	2760
	AGATGACGCA	AGATGACGCA	AGATGACGCA	AGATGACGCA	AGATGACGCA	AGATGACGCA	2820
	AGATGACGCA	AGATGACGCA	AGATGACGCA	AGATGACGCA	AGATGACGCA	AGATGACGCA	2880
	AGATGACGCA	AGATGACGCA	AGATGACGCA	AGATGACGCA	AGATGACGCA	AGATGACGCA	2940
50	AGATGACGCA	AGATGACGCA	AGATGACGCA	AGATGACGCA	AGATGACGCA	AGATGACGCA	3000
	AGATGACGCA	AGATGACGCA	AGATGACGCA	AGATGACGCA	AGATGACGCA	AGATGACGCA	3060
	AGATGACGCA	AGATGACGCA	AGATGACGCA	AGATGACGCA	AGATGACGCA	AGATGACGCA	3120

80      Seq ID NO: 48 Protein sequence  
Protein Accession #: NP\_066008.1

WO 02/098358

PCT/US02/17594

1  
MEICKMIDNR | QRQVLVFFVL | LSLSCAGL | GYSVVEHTE | RGSFVANKG | DLGLGLTMS 60  
TRKARLISQK | NQKHQLKQK | TQDLINEXL | DESELOPTE | FCIHLFVLM | ENPLSTPOAR 120  
LEVIDINDHS | MPWTEKIML | KIPENSLPT | EPLNANALG | DVGNHVNQY | KIGPSHPRV 180  
LTHPEEDRKH | YHELVLKSL | DBREBQJRL | PLTALQSGP | PRGTAQAVI | EVDLINDAP 240  
EFQPIYVQ | FHSPLFSL | VATUSHLLD | GSNKILST | LQPSBDSK | TLEVWPTGR 300  
VLEKQVDFE | MYSYVEVRLK | ATQOGLSKK | CTLLQVVDV | NENPQVIMS | ALTSPISNS 360  
PELVAVFV | SDPBGSGNK | TISSIGBPL | FLKPSKMF | YLIVTERALD | REARAVNIT 420  
LUTVMOTPR | LSTHETVQ | LSEVBNADK | PFTSTYPLP | NBSBEPALI | GSVNATDRI 480  
GTNMQVYSL | LEPQLPRL | ASLVBNADK | CLPLALSLD | VYALSEFE | VENTATREDA 540  
LSBLAVLVI | VLDANESPF | VYLPQLQNSA | PCTELVFRP | ECTVLTUV | AVDGSGQNA 600  
WLSYOLLAT | EPLCPGVWH | NUBVETAREL | SERDANRGL | VLVVDNGEP | PRSATATLW 660  
LVVQDSFSP | LPLPEAAPQ | TQNSLTVL | VVALSVSL | FLPSVLLFA | VRLKESRAA 720  
SVORCSFSG | PPSGLVWVS | CTTLGLQSG | YEVCLQSSS | TSPFLKPI | ITPSP

Seq ID NO: 49 DNA sequence  
Nucleic Acid Accession #: CAT cluster

1  
TTTTTTTTT | ATAATACACA | GACTTTAAT | AAATGTGAC | TAAATATAA | TGTCTAATA 60  
ATAATGAAT | CTGATCTTA | CATCTTAAT | TACTTTTAA | TCTTTTAT | GTGCTATT 120  
TCTCTAGGT | TCTCTTGTCT | TAGTTTGT | AAAGTTCCT | ATTTTTGA | TAATGTAGA 180  
TATCATTAAT | AAGGAAATAT | CAGGAAATAT | AATATGAAG | AGAAACAT | AGCTATGTC 240  
AACAATAAP | AATATGCGA | ACTCTTAAGC | ACATGAATC | TGTATTATT | GTACAGCAT 300  
CAATANNTT | ACTCTTCA | TGTCTAAGT | GAGCTAGAC | CTGACAGA | CTGACAGA 360  
TAGGAAATA | AGGAATTTT | CCAACATAT | TAATATTAT | GAAATGTG | AACTTACAG 420  
TTAATACATA | AGTATGAA | AATGATGAT | TTTAAGGA | TCTGAATA | TTA

Seq ID NO: 50 DNA sequence  
Nucleic Acid Accession #: AF034799.1  
Coding sequence: 170..3943

1  
CATTCGGGA | CGCAAGTGG | GACAGAGAT | GGTGTAGCT | CTTACCGGC | TGCAGCAGG 60  
GAATAGGTC | AGAGATGCG | GGTGTAGCT | TGTCTTCTC | TACAGCGGA | GATCTGCTT 120  
GTCTTAGAT | ACTTAAGAGA | CACAGAGAT | GAATCTAGCA | TGTCTTCTC | TGTCTTCTC 180  
AGTGAATCT | AGATATGCG | AGAGATGCG | AATGAGTCT | AAGAGATG | AAGAGATG 240  
CTCAGATCT | GATCTGCTT | TGTCTTCTC | GATCTGCTT | AATGAGTCT | AAGAGATG 300  
TGTCTTCTC | AGCTCTGCG | AGAGATGCG | AGCTCTGCG | CTGAGTCT | AAGAGATG 360  
GATCTGCTT | TGTCTTCTC | ACTCTTCTC | GATCTGCTT | AATGAGTCT | AAGAGATG 420  
TGTCTTCTC | TGTCTTCTC | GATCTGCTT | TGTCTTCTC | GATCTGCTT | AAGAGATG 480  
TGTCTTCTC | AAGAGATG | AAGAGATG | AAGAGATG | AAGAGATG | AAGAGATG 540  
CTCTGAGCT | AAGAGATG | AAGAGATG | AAGAGATG | AAGAGATG | AAGAGATG 600  
TGTCTTCTC | AAGAGATG | AAGAGATG | AAGAGATG | AAGAGATG | AAGAGATG 660  
CTCTGAGCT | TGTCTTCTC | AAGAGATG | AAGAGATG | AAGAGATG | AAGAGATG 720  
CTCTGAGCT | TGTCTTCTC | AAGAGATG | AAGAGATG | AAGAGATG | AAGAGATG 780  
ACTGAGGGA | GATCTGCTT | GATCTGCTT | GATCTGCTT | GATCTGCTT | GATCTGCTT 840  
TGTCTTCTC | AAGAGATG | AAGAGATG | AAGAGATG | AAGAGATG | AAGAGATG 900  
GATCTGCTT | GATCTGCTT | GATCTGCTT | GATCTGCTT | GATCTGCTT | GATCTGCTT 960  
GATCTGCTT | AAGAGATG | AAGAGATG | AAGAGATG | AAGAGATG | AAGAGATG 1020  
GATCTGCTT | AAGAGATG | AAGAGATG | AAGAGATG | AAGAGATG | AAGAGATG 1080  
GATCTGCTT | AAGAGATG | AAGAGATG | AAGAGATG | AAGAGATG | AAGAGATG 1140  
GATCTGCTT | AAGAGATG | AAGAGATG | AAGAGATG | AAGAGATG | AAGAGATG 1200  
GATCTGCTT | AAGAGATG | AAGAGATG | AAGAGATG | AAGAGATG | AAGAGATG 1260  
GATCTGCTT | AAGAGATG | AAGAGATG | AAGAGATG | AAGAGATG | AAGAGATG 1320  
GATCTGCTT | AAGAGATG | AAGAGATG | AAGAGATG | AAGAGATG | AAGAGATG 1380  
GATCTGCTT | AAGAGATG | AAGAGATG | AAGAGATG | AAGAGATG | AAGAGATG 1440  
GATCTGCTT | AAGAGATG | AAGAGATG | AAGAGATG | AAGAGATG | AAGAGATG 1500  
GATCTGCTT | AAGAGATG | AAGAGATG | AAGAGATG | AAGAGATG | AAGAGATG 1560  
GATCTGCTT | AAGAGATG | AAGAGATG | AAGAGATG | AAGAGATG | AAGAGATG 1620  
GATCTGCTT | AAGAGATG | AAGAGATG | AAGAGATG | AAGAGATG | AAGAGATG 1680  
GATCTGCTT | AAGAGATG | AAGAGATG | AAGAGATG | AAGAGATG | AAGAGATG 1740  
GATCTGCTT | AAGAGATG | AAGAGATG | AAGAGATG | AAGAGATG | AAGAGATG 1800  
GATCTGCTT | AAGAGATG | AAGAGATG | AAGAGATG | AAGAGATG | AAGAGATG 1860  
GATCTGCTT | AAGAGATG | AAGAGATG | AAGAGATG | AAGAGATG | AAGAGATG 1920  
GATCTGCTT | AAGAGATG | AAGAGATG | AAGAGATG | AAGAGATG | AAGAGATG 1980  
GATCTGCTT | AAGAGATG | AAGAGATG | AAGAGATG | AAGAGATG | AAGAGATG 2040  
GATCTGCTT | AAGAGATG | AAGAGATG | AAGAGATG | AAGAGATG | AAGAGATG 2100  
GATCTGCTT | AAGAGATG | AAGAGATG | AAGAGATG | AAGAGATG | AAGAGATG 2160  
GATCTGCTT | AAGAGATG | AAGAGATG | AAGAGATG | AAGAGATG | AAGAGATG 2220  
GATCTGCTT | AAGAGATG | AAGAGATG | AAGAGATG | AAGAGATG | AAGAGATG 2280  
GATCTGCTT | AAGAGATG | AAGAGATG | AAGAGATG | AAGAGATG | AAGAGATG 2340  
GATCTGCTT | AAGAGATG | AAGAGATG | AAGAGATG | AAGAGATG | AAGAGATG 2400  
GATCTGCTT | AAGAGATG | AAGAGATG | AAGAGATG | AAGAGATG | AAGAGATG 2460  
GATCTGCTT | AAGAGATG | AAGAGATG | AAGAGATG | AAGAGATG | AAGAGATG 2520  
GATCTGCTT | AAGAGATG | AAGAGATG | AAGAGATG | AAGAGATG | AAGAGATG 2580  
GATCTGCTT | AAGAGATG | AAGAGATG | AAGAGATG | AAGAGATG | AAGAGATG 2640  
GATCTGCTT | AAGAGATG | AAGAGATG | AAGAGATG | AAGAGATG | AAGAGATG 2700  
GATCTGCTT | AAGAGATG | AAGAGATG | AAGAGATG | AAGAGATG | AAGAGATG 2760  
GATCTGCTT | AAGAGATG | AAGAGATG | AAGAGATG | AAGAGATG | AAGAGATG 2820  
GATCTGCTT | AAGAGATG | AAGAGATG | AAGAGATG | AAGAGATG | AAGAGATG 2880

WO 02/098358

PCT/US02/17594

5  
10  
15  
20

CGCATGCTA GAGCTTTGT TGGGAATGC TCGTGGTAC GTGCGAGCT GCGGAGCAA 2940  
CTGAGAGAT GTGGCATCA TGTCTGCTT ATCTGACACT GAGATCCGA GAGAAATGG 3000  
AATCGCAAT CCACTGCATC GCTTAAACT TCGATTAGA ATCCAGGAG TGGTTTCCCT 3060  
AACAGTCTC TCGAGTCTC CACATCTCG AACTCTTTCA GCGACGTTT GGTGTACTCA 3120  
TGAAGAAAT GAAATCTTG CACTTGCAC AAGAAAGAA GAATCTGAG AAGGAAGCTG 3180  
GGCCAGTGT CCGGTTTTT TACAGACCT GCTTATGGA GATATGATC ATGAGTGGAT 3240  
TGGAAATGA TGGCTTCCA GCTTGGGTT AACTGATAC AGAATTACT TTATGGAATG 3300  
CTTGATGAT GCAAGAAAT TATGATACCT AACAAAAAA GATCTCCGT TCAATTAAAT 3360  
AATGTGATC AGTTTCCAT GACAAATTT ACATATGGA ATATATGCT TAAAGAGTT 3420  
GAATTATAC AGAAGAAAC TGAAGAGAG ACGGGAAGA AGCCAACTG AATATAAGA 3480  
CGTTTGGTG TGGGCAATG ACCGAGTAT TCGCTGGTA CAGCAATTG GACTTCAGA 3540  
ATATGCAAT AATATACTG AGAAGCGGT GATGCTGCTA CTATAGACC TGGATGAAA 3600  
CTTTCATAC AGCACTTAG CTATTATAT AGCAATCTCA ACAGAGACA CCGAGCAAG 3660  
CGAGTCTCT GAGAGGATC ACATTAACCT CTTCGGCTTG GCACTGAAA GCGCACTGA 3720  
TGAAGTAT GAGAGAACT TCGAGCTTG ATCAACTTG AGAAGCAAT TTCCTCTTG 3780  
TGAAGTAT GAGATGAGA TATGCTCTG GTCTTCAGA ACATTAACG CTGGATTAG 3840  
GTAAACCA ACCCTCTGG AGTCAAGAA AATGAGACA GATGTTCTT CATTAAGAT 3900  
GCGAGTTA GCACTCTCA CTTCTGCAC ATCTCTCTT TACAGACA CTCAGAGG 3960  
CGACACTGA CCGCTATGG CTTCTTTCA GCTATCTCA CTTAAGTGC ACTACATCT 4020  
AAGAGAGA CAGTGAAGA CTTTGTGAA AACTGAATC

Seq ID NO: 51 Protein sequence  
Protein Accession #: AAC2100.1

25  
30  
35  
40  
45

1 11 21 31 41 51  
| | | | |  
NNECVNPTIN EDTNPGQRGS QSSGSDSDSH FGLAVNMLD EEDRLDLTL ETQESLSLAQ 60  
QLQGVYDTR DSIQRQLNSA LPQDSSLTG GLAGSKGADP PEPALITGL NACREQLLEK 120  
BSISSELNAS DNTRELLLEH LSCLVNKHNR SLAMTVVKRQ AQSPGVSSSE VEVLLALESL 180  
FERRALDEK VRELEVLSE KYMALSELA AAGQETVALA DQNVITGRWV ASSGPTSE 240  
HLESGEPOOK VHEKRLNSG IDSDITSGI VELQELLEKO NYEMAKIKER LAALSSRVGE 300  
VQBAETARK DLIKEENMT KYRDIRAM AQKEMBBRI TTLERKYLGA QRETSIDIN 360  
NDKLENKAN KPAILRQNE KIKQLQSEL LABEKLQTM RPAETLPVE ASLACRIAL 420  
TASETHONI ERMRELSQ LEEIQMLGR AQREDSHSE RKKLDTVD LLLTDSNELL 480  
QHLKESNA LERGVLTG STSPRLSE SLADKSLAE EIKGLSELD QJMECTSLI 540  
EPTIKTHLD TAEKLYSVG SLVDSQDYR TKVIRRFRR GEMVREDBP KVELSDHSM 600  
NRTQIQVLG SHPFSDTSM SDIDDDRET IFESDLSGP SORSDAQTLA MMLQQLDAI 660  
HKEILQZER KSTELRAE ERYVASVL SGLSLAWHF GELSTAVTA SLSASSPFS 720  
GISTPILTR SPANCRMG VNLPSGLK VIKLAVVE DQREDAKIK CTSPPTTZR 780  
ALRWITLPS SYNDARSEL SVLEPBSIG LGSANSQDS LHKAPKXGI KSIIGRLFK 840  
KEKARLOLR QPWTSAARQ ESLGLKLET QAEDRELEK KHSLLEARR KGLFFAQWDG 900  
PTVYVLELM LQNPATYAA CHANYSGLAI KSLSDTEIG REIGTSLPL KLLALALQ 960  
NVLSGSDAP STSTGSGV NYTEEMENL AAKATYKSE DSGACQVY LQTLAYDMM 1020  
NEWICNWLK SLGLQYSRY FVECLVDAM LRLTKKDLR VELAWVDSH RTSLOVGM 1080  
LKLNYDREK LBRREBARG EIKDVLVNH DRVIRWIAL QRETAANNIL BSOVHSLIA 1140  
LDENFDYSL ALLQLPTQ TQARQLRES DNMILALGT RELESDDM FREGSTNRQ 1200  
FFPREVHGS MUPGSESLP AGFALTTGQ GSKMTTVA SRLQLQNS TVRTYGC



It is understood that the examples described above in no way serve to limit the true scope of this invention, but rather are presented for illustrative purposes. All publications, sequences of accession numbers, and patent applications cited in this specification are herein  
5 incorporated by reference as if each individual publication or patent application were specifically and individually indicated to be incorporated by reference.

WHAT IS CLAIMED IS:

1. A method of detecting an androgen-independent prostate cancer cell in a sample from a patient having undergone androgen ablation therapy, the method comprising determining the presence or absence of a nucleic acid comprising a sequence at least 80% identical to a sequence as shown in Tables 1A-4.
2. The method of claim 1, wherein said determining is by hybridizing with a polynucleotide that selectively hybridizes to a sequence at least 95% identical to a sequence as shown in Tables 1A-4.
3. The method of claim 1, wherein the biological sample:
  - a) is a tissue sample; or
  - b) comprises isolated nucleic acids.
4. The method of claim 3:
  - a) wherein the nucleic acids are mRNA; or
  - b) further comprising the step of amplifying nucleic acids before the step of contacting the biological sample with the polynucleotide.
5. The method of claim 2, wherein the polynucleotide:
  - a) comprises a sequence as shown in Tables 1A-4;
  - b) is labeled, including a fluorescent label; or
  - c) is immobilized on a solid surface.
6. The method according to claim 1, wherein said biological sample is contacted with a plurality of polynucleotides that each selectively hybridizes to a sequence at least 95% identical to a first sequence as shown in Tables 1A-4.
7. The method according to claim 6, wherein said plurality of polynucleotides are immobilized on a solid surface.

8. An isolated polypeptide which is encoded by a nucleic acid molecule having polynucleotide sequence as shown in Tables 1A-4.
9. An antibody that specifically binds a polypeptide of claim 8.
10. The antibody of claim 9:
  - a) further conjugated to an effector component, including a fluorescent label a radioisotope or a cytotoxic chemical; or
  - b) which is an antibody fragment or humanized antibody.
11. A method of detecting an androgen-independent prostate cancer cell in a patient having undergone androgen ablation therapy, the method comprising contacting a sample from said patient with an antibody of claim 9.
12. The method of claim 11, wherein:
  - a) the antibody is further conjugated to an effector component, e.g., a fluorescent label; or
  - b) said sample comprises a cell.
13. A method of detecting antibodies specific to androgen-independent prostate cancer in a patient having undergone androgen ablation, the method comprising contacting a biological sample from the patient with a polypeptide encoded by a nucleic acid comprising a sequence from Tables 1A-4.
14. A method of inhibiting proliferation of androgen-independent prostate cancer cells in a patient having undergone androgen ablation therapy, the method comprising administering to the patient a therapeutically effective amount of a compound that specifically eliminates cells expressing an antigen listed in Tables 1A-4.
15. The method of claim 14, wherein the compound is an antibody.
16. A drug screening assay comprising the steps of:

- a) administering a test compound to a mammal having a prostate proliferative condition or a cell isolated therefrom;
  - 65 b) comparing the level of gene expression of a polynucleotide that selectively hybridizes to a sequence at least 80% identical to a sequence as shown in Tables 1A-4 in a treated cell or mammal with the level of gene expression of the polynucleotide in a control cell or mammal, wherein a test compound that modulates the level of expression of the polynucleotide is a candidate for the  
70 treatment of prostate cancer.
17. The assay of claim 16, wherein:
- a) the control is a mammal with prostate cancer or a cell therefrom that has not been treated with the test compound; or
  - 75 b) the control is a normal cell or mammal.
18. A method for treating a mammal having a prostate proliferative condition or prostate cancer comprising administering a compound identified by the assay of claim 16.
- 80 19. A pharmaceutical composition for treating a mammal having a prostate proliferative condition or prostate cancer, the composition comprising a compound identified by the assay of claim 16 and a physiologically acceptable excipient.
20. A method of detecting a prostate cancer associated transcript, the method comprising  
85 contacting a biological sample from the patient with a plurality of polynucleotides wherein at least two of said polynucleotides selectively hybridize to a difference sequence at least 80% identical to a sequence as shown in Tables 1A-4.
21. A method of detecting a prostate cancer, the method comprising the steps of:
- 90 a) providing a biological sample from a patient;
  - b) contacting the biological sample with a first polynucleotide that selectively hybridizes to a sequence at least 80% identical to a first sequence as shown in Tables 1A-4, to determine the level of a prostate cancer-associated transcript in the biological sample; and with a second polynucleotide that selectively

- 95                    hybridizes to a second sequence at least 80% identical to a sequence not  
shown in Tables 1A-4; wherein the expression of said second sequence is not  
substantially changed in prostate cancer, to determine the level of expression  
of a control transcript in the biological sample; and  
c) comparing the level of the prostate cancer-associated transcript to a level of the  
100                    normal tissue associated transcript in the biological sample.

22.     A method for quantitation of a prostate cancer-associated transcript in a cell from a  
patient, the method comprising contacting a biological sample from the patient with a  
polynucleotide that selectively hybridizes to a sequence at least 80% identical to a sequence  
105     as shown in Tables 1A-4.

23.     The method of claim 22, wherein:
- a) the polynucleotide selectively hybridizes to a sequence at least 95% identical to a  
sequence as shown in Tables 1A-4;
  - 110     b) the biological sample is a tissue sample;
  - c) the biological sample comprises isolated nucleic acids;
  - d) the nucleic acids are mRNA;
  - e) further comprising the step of amplifying nucleic acids before the step of  
contacting the biological sample with the polynucleotide;
  - 115     f) the polynucleotide comprises a sequence as shown in Tables 1A-4;
  - g) the polynucleotide is labeled, including a fluorescent label; or
  - h) the polynucleotide is immobilized on a solid surface.

24.     A biochip comprising a plurality of polynucleotides that selectively hybridize to a  
120     sequence at least 80% identical to a sequence as shown in Tables 1A-4.

25.     A method of screening drug candidates comprising:
- a) providing a cell that expresses an expression profile gene selected from the group  
consisting of an expression profile gene set forth in Tables 1A-4 or fragment  
125     thereof;
  - b) adding a drug candidate to said cell; and

WO 02/098358

PCT/US02/17594

- c) determining the effect of said drug candidate on the expression of said expression profile gene.

- 130 26. A method according to claim 22 wherein said determining comprises comparing the level of expression in the absence of said drug candidate to the level of expression in the presence of said drug candidate.